Location based Network Optimizations for Mobile Wireless Networks

A study of the Impact of Mobility and Inaccurate Information

Nielsen, Jimmy Jessen

Publication date:
2011

Document Version
Accepted author manuscript, peer reviewed version

LOCATION BASED NETWORK OPTIMIZATIONS FOR MOBILE WIRELESS NETWORKS

- A study of the Impact of Mobility and Inaccurate Information -

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF
ELECTRICAL ENGINEERING
OF
AALBORG UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Jimmy Jessen Nielsen
March 24, 2011

Supervisors:
Assoc. Prof. Hans-Peter Schwefel, Aalborg University, Denmark
Assoc. Prof. Tatiana Kozlova Madsen, Aalborg University, Denmark

The assessment committee:
Prof. Giuseppe Bianchi, University of Roma Tor Vergata, Italy
Assoc. Prof. Atílio Gameiro, Instituto de Telecomunicações, Portugal
Assoc. Prof. Troels B. Sørensen, Aalborg University, Denmark (Chairman)

Moderator:
Prof. Bernard Henri Fleury, Aalborg University, Denmark

Date of defence: May 10, 2011

Report: R03-2011
ISSN: 0908-1224
ISBN: 978-87-92328-66-3

Dedicated to my life partner, Jill.

Abstract

With an expected annual doubling in mobile data volumes, wireless operators are likely to face a major challenge in renewing and scaling their network infrastructure to support the future demands. Since cellular technologies might not be able to satisfy this demand single-handedly, local connectivity options such as Wi-Fi hotspots are seen as a means for relieving the load on the cellular networks. Smart phones and laptop/netbook computers are typically equipped with both cellular and Wi-Fi connectivity options, which allows the user or terminal to connect to Wi-Fi access points when available.

In addition to cellular and Wi-Fi connectivity, smart phones are typically also equipped with Global Positioning System (GPS) receivers that enable outdoor localization. Further, cooperative network-based localization algorithms that enable indoor localization are gaining interest in the research community. This motivates the main topic of this thesis, which is to combine the connectivity and localization topics and investigate how useful location based network optimizations are for improving the user's last hop connectivity. The main focus is on analyzing the impact of inaccurate location information, user mobility, measurement collection delays, and model parameter inaccuracies.

The first contribution in this thesis concerns the impact of realistic communication constraints on network-based localization. In this joint work, the contribution is the modeling of the message exchanges necessary for centralized, cooperative network-based localization, as well as imposing realistic network constraints on these message flows. These constraints introduce additional delays in the message exchanges, which have a significant impact on tracking performance of the localization algorithm compared to the ideal situation where measurements are assumed to be instantaneously available.

Secondly, two-hop relaying is considered for improving throughput and outage performance of downlink transmissions from a Wi-Fi hotspot to mobile destination nodes. The main focus is on relay selection, i.e. determining whether to use direct or relayed transmission and in the latter case, which mobile relay node to use. Here, two schemes are investigated: the first uses collected SNR measurements from all links, whereas the second uses collected position measurements. The performance of the schemes is investigated in

terms of throughput performance in case of inaccurate input parameters, such as inaccurate position information, missing or crude Non-Line Of Sight (NLOS) information, and delayed measurements. Simulations have been used to investigate the performance of the schemes in different cases and the impact of different sources of inaccuracies.

Thirdly, an extension to the two-hop relaying case is considered. This extension is a model-based cross-layer optimization, which uses position information to predict intra-system interference between different relay-destination pairs and to adapt relay transmit power levels, which thereby enables simultaneous relay-destination transmissions. Simulation results have shown that a significant throughput improvement is possible using this technique. In order to estimate the achievable throughput performance of different configuration choices, a distance dependent throughput model was developed.

Fourthly, location based algorithms for prediction and planning of handovers between cellular and Wi-Fi networks is considered. The thesis focuses on the single user case, where the mobile user's device can connect to a cellular network or one of many available Wi-Fi networks. A priori information regarding the average achievable throughput for the different networks is assumed to be available. An algorithm for determining the sequence of handovers that leads to the maximum throughput within a fixed time horizon is proposed. Also, a more efficient heuristic algorithm is proposed, and the performance of these algorithms is compared to a simple hysteresis-based algorithm, an upper bound reference algorithm, and a static algorithm that is always connected using the cellular network.

In summary, this thesis has investigated the usefulness of location information for last hop route selection. The results have shown that while location-based two-hop relaying can help to decrease the signaling overhead and cope well with mobility, it requires a high-accuracy localization system as well as approximate environmental knowledge in order to cope with NLOS situations. It is also shown how location information enables simultaneous relay-destination transmissions through interference-awareness, which can further increase throughput performance. Finally, location-based handover prediction is shown to work well if the user's movement trajectory is accurately estimated. The proposed heuristic algorithm is feasible in terms of computational complexity, and its performance is comparable to the optimal solution.

Dansk Resumé

Med en forventet årlig fordobling af datatraffik har mobilnetværksoperatørerne en stor udfordring i sigte med at forny og opgradere deres netværksinfrastrukturer for at kunne følge med fremtidige behov. Da mobilnetværkene ikke nødvendigvis kan opfylde dette behov for datatraffik alene, kan lokale forbindelser såsom Wi-Fi adgangspunkter udnyttes som aflastning af det mobile netværk. Idet smart phones og bærbare computere typisk er udstyret med forbindelsesteknologier til både mobilnetværk og Wi-Fi netværk, er det muligt for brugeren at skifte til en tilgængelig lokal Wi-Fi forbindelse.

Udover forskellige forbindelsesmuligheder, har smart phones typisk også en Global Positioning System (GPS) modtager indbygget, hvilket muliggør udendørs lokalisering. Desuden er der en stigende interesse for kooperative, netværksbaserede lokaliseringsalgoritmer i forskningsmiljøet, der muliggør højpræcisions lokalisering indendøre. Dette motiverer hovedtemaet i denne afhandling, som går ud på at kombinere netværksforbindelses- og lokaliseringsemnerne, og undersøge hvor brugbare lokationsbaserede netværksoptimeringer er til at forbedre brugerens sidste-hops netværksforbindelse. Hovedfokus er på at analysere påvirkningen af unøjagtigheder i lokationsinformation, brugerens bevægelser, indsamling af målinger og modelparameterunøjagtigheder.

Det første bidrag i afhandlingen omhandler betydningen af en realistisk model af kommunikationen til netværksbaseret lokalisering. I dette samarbejde består bidraget af en model af beskedudvekslingen som er nødvendig for centraliseret, kooperativ lokalisering, såvel som indførslen af realistiske begrænsninger på denne kommunikation. Disse begrænsninger medfører forsinkelser i beskedudvekslingerne, som derved har en væsentlig betydning for ydelsen af de betragtede lokaliseringsalgoritmer til sammenligning med den ideelle situation, hvor det antages at målinger er øjeblikkeligt tilgængelige.

Dernæst betragtes to-hop radio-relæer som en metode til at forbedre overførselshastigheden og reducere udfald for datatransmissioner fra et Wi-Fi adgangspunkt til en bevægelig destinationsenhed. Hovedfokus er på udvælgelsen af relæer, dvs. at afgøre om en transmission skal foregå direkte eller via et relæ og i så fald hvilket relæ der skal anvendes. Her undersøges to metoder: den første benytter indsamlede SNR målinger fra alle forbindelser, hvorimod

den anden metode anvender indsamlet lokationsinformation. Ydelsen af de to metoder er blevet undersøgt i forhold til den opnåelige overførselshastighed i tilfælde af unøjagtige input-parametre såsom lokationsinformation, utilstrækkelig information om forhindringer og forsinkede målinger. Simuleringer er blevet anvendt til at undersøge ydelsen af de forskellige metoder i disse situationer.

Det tredje bidrag består af en udvidelse af den ovennævnte metode baseret på radio-relæer. Denne udvidelse er en optimering på tværs af protokollag, som bruger lokationsinformation til at forudsige indbyrdes interferens mellem forskellige relæ-destinationspar samt til at vælge relæer og at indstille relæernes transmissionsstyrke. Dette muliggør samtidige relæ til destinationstransmissioner. Simuleringsresultater har vist at denne teknik giver signifikante forbedringer af overførselshastigheden. Til at estimere den mulige overførselshastighed af forskellige mulige konfigurationer, blev der udviklet en afstandsafhængig model af overførselshastigheden.

Fjerde og sidste bidrag er en lokationsbaseret algorithme til forudsigelse og planlægning af overdragelse af netværksforbindelser i og imellem mobilnetværk og lokale Wi-Fi netværk. Dette arbejde tager udgangspunkt i en enkelt bruger, hvor denne mobile brugers netværksenhed er i stand til at forbinde til mobilnetværket eller ét af mange tilgængelige Wi-Fi netværk. Forhåndsinformation om den forventet opnåelige overførselshastighed for de forskellige netværk antages at være tilgængelig. På baggrund af dette er der blevet udviklet en algoritme til at finde den sekvens af netværksoverdragelser indenfor en given tidshorisont, som resulterer i den højeste gennemsnitsoverførselshastighed. Derudover foreslås også en mere effektiv heuristisk algoritme, og ydelsen af disse algoritmer sammenlignes med en simpel hysterese-baseret algoritme, en reference-algoritme, der giver en øvre grænse samt en statisk algoritme for tilfældet hvor kun det mobile netværk bruges.

I korte træk viser denne afhandling en undersøgelse af brugbarheden af lokationsinformation for valg af sidste-hops netværksforbindelse. Resultater har vist at mens lokationsbaserede to-hops relæ-teknikker giver et lavere signalerings-overhead og håndterer mobilitet godt, kræves der også et højpræcisions lokaliseringssystem samt anslået viden om det lokale miljø for at kunne håndtere f.eks. udsynsforhindringer. Det er også blevet vist hvordan lokationsinformation kan muliggøre samtidige relæ-destinationstransmissioner gennem interferens-bevidsthed, hvilket yderligere forbedrer ydelsen. Endelig er det blevet vist at lokationsbaseret netværksoverdragelse virker godt såfremt brugerens bevægelser kan forudsiges præcist. Den foreslåede heuristiske algoritme er beregningsmæssig let nok til at kunne afvikles online og dens ydelse er næsten på højde med den optimal algoritme.

Acknowledgements

Thanks to my supervisor Hans-Peter Schwefel for always managing to squeeze in 5-10 minutes for discussions, no matter how busy you are, and for our many interesting discussions and brainstorming sessions, which have helped to shape my PhD.

It has also been a great pleasure to have Tatiana K. Madsen as my second supervisor because she has always had an open door and been ready to discuss and give me a different perspective on my day-to-day work.

Further, I would like to thank my colleagues in NetSec and other friends at Aalborg University, who have contributed to a friendly and joyful working environment.

Also, a big thanks for all my colleagues in the WHERE project, who have given me valuable inputs and new perspectives on my work.

Finally I owe a big thanks to my life's partner Jill and my family for their unconditional love and support.

WHERE Project Acknowledgement

This work has been performed in the framework of the ICT projects ICT-217033 WHERE and ICT-248894 WHERE2, which are partly funded by the Framework Program 7 of the European Commission.

Contents

List of Acronyms **xi**

1 Introduction **1**
1.1 Motivation and Main Problem 2
1.2 Scenarios 5
1.3 Network Architecture 8
1.4 Objectives 10
1.5 Methodology 12
1.6 Contributions 12

2 Background **19**
2.1 Wi-Fi Performance Aspects 20
2.2 Relaying 24
2.3 Network-based Localization 26
2.4 Mobility Models 32

3 Distance-dependent Throughput Model for WiFi **41**
3.1 Introduction 42
3.2 Related Work 44
3.3 Throughput Model 45
3.4 Results and Discussion 48
3.5 Conclusion 51

4 Realistic Communication Constraints for Localization **54**
4.1 Introduction 55
4.2 System Model 56
4.3 Measurement Collection 57
4.4 Localization Algorithms 62
4.5 Evaluation Methodology 63
4.6 Results and Discussion 65
4.7 Conclusion 71

5 Location based Relay Selection **75**
5.1 Introduction 76
5.2 Related Work 78
5.3 Centralized Relay Selection 80
5.4 IEEE 802.11 based Evaluation Scenario 84
5.5 Evaluation Methodology 86
5.6 Results and Discussion 91
5.7 Conclusion 106

6 Joint Location based Relay Selection and Power Adaptation **110**
6.1 Introduction 111
6.2 Related Work 113
6.3 The SimTX Scheme 115
6.4 Throughput Calculation 118
6.5 Evaluation Methodology 119
6.6 Results and Discussion 121
6.7 Conclusion 126

7 Location based Handover Optimization **131**
7.1 Introduction 132
7.2 Related Work 135
7.3 System Model 138
7.4 Look-ahead Prediction Algorithms 141
7.5 Reference Algorithms 144
7.6 Evaluation Methodology 145
7.7 Results and Discussion 150
7.8 Implementation Considerations 159
7.9 Conclusion and Outlook 160

8 Conclusions and Outlook **165**

A Timing Specifications of IEEE 802.11a **173**

List of Acronyms

3GPP 3rd Generation Partnership Project

AF Amplify-Forward

AN Anchor Node

AOA Angle of Arrival

AP Access Point

BER Bit Error Rate

BSS Basic Service Set

cdf Cumulative Distribution Function

CF Compress-Forward

CP Cooperative Positioning

CSMA/CA Carrier Sense Multiple Access with Collision Avoidance

DCF Distributed Coordination Function

DF Decode-Forward

DHCP Dynamic Host Configuration Protocol

DIFS Distributed InterFrame Space

EKF Extended Kalman Filter

ESS Extended Service Set

FER Frame Error Rate

GN Gauss-Newton

GNSS Global Navigation Satellite System

GPS Global Positioning System

HHO horizontal handover

HO handover

HSDPA High-Speed Downlink Packet Access

HSPA High-Speed Packet Access

i.i.d. independent and identically distributed

IP Internet Protocol

LOS Line-of-Sight

LTE Long-Term Evolution

LS Least Squares

MAC Medium Access Control

MICS Media-Independent Command Service

MIES Media-Independent Event Service

MIH Media Independent Handover

MIHF MIH Functions

MIIS MIH Information Server

MIP Mobile IP

mmPr Mismatch Probability

MS Mobile Station

MTU Maximum Transmission Unit

NLOS Non-Line Of Sight

P2P Peer-to-Peer

PCA Principal Component Analysis

PCF Point Coordination Function

POMDP Partially Observable Markov Decision Process

QoS Quality of Service

RSS Received Signal Strength

RTS/CTS Request to Send / Clear to Send

RV Random Variable

SIFS Short InterFrame Space

SNR Signal to Noise Ratio

SINR Signal to Interference and Noise Ratio

TDOA Time Difference of Arrival

TLS Total Least Squares

TOA Time of Arrival

UWB ultra-wideband

VHO vertical handover

WAAN Wide-Area Access Network

WAN Wide-Area Network

WCDMA Wideband Code Division Multiple Access

1

Introduction

Initially, the main research problem of investigating how well location information can be used for network optimizations is introduced. This problem is considered for the example use case of cellular offloading, specifically for handover and relaying techniques, which are motivated from large and small scale scenarios. Further the specific objectives of this work, which include studying the impact of user mobility, collection delays, and inaccurate model parameters, are defined. Finally a summary of the specific contributions and the corresponding publications is given.

1.1 Motivation and Main Problem

Today's mobile devices such as smart phones, tablets, netbooks and laptop computers are typically equipped with multiple radio interfaces, which make them able to connect to cellular 3G/4G networks and local Wi-Fi networks. Besides having both cellular and Wi-Fi connectivity options, smart phones are typically also equipped with a GPS receiver, which can provide the location of the device with only a few meters of inaccuracy in outdoor scenarios where the GPS receiver has a clear view to the sky.

The availability of location information has led to the emergence of many location based services. Examples include car navigation, intelligent traffic systems, store finder, location based advertisement, and augmented reality applications. Such services and applications have requirements to the position accuracy in the order of tens of meters down to or even below a few meters in both outdoor and indoor scenarios. One challenge is however that in indoor scenarios or urban canyons, where the GPS receiver does not have a clear view to at least a handful of GPS satellites in the sky, localization via GPS can often not provide the user with his or her location or the accuracy is insufficient.

This deficiency of traditional GPS-based systems has motivated an interest in the research community for localization solutions that are able to exploit the available wireless networks in mobile devices for localization. In the research community, such localization solutions that exploit other technologies as well as cooperation between mobile devices to provide high accuracy and coverage are being researched, for instance in the scope of the EU-funded research project WHERE and its continuation WHERE2. In addition to providing accurate location information, an emerging topic is the exploitation of location information for network optimizations.

Growing Data Volumes

As mentioned above, mobile devices are able to connect to both cellular 3G/4G networks and local Wi-Fi networks, thereby delivering ubiquitous connectivity to online services such as email, browsing, social networking, download/streaming of music or video, and others, where many of those can be quite data-consuming. Also, since the market penetration of these mobile devices increases and more users start taking advantage of the devices' various applications and services, the amount of mobile data being served by the mobile operators' networks is expected to grow quickly. According to [Cisco

Systems, 2010], Cisco forecasts an increase in mobile data of 108% per year in the period $2009 - 2014$.

If this future demand for mobile data needs to be satisfied primarily through the mobile operators' cellular networks, it is questionable if the mobile operators are able to scale their infrastructure and renew their technologies quickly enough to keep up with the constantly rising demand. The biggest bottleneck is here the last hop wireless link, since it depends on the deployment of base stations and the capacity of the transmission technologies such as High-Speed Packet Access (HSPA) and future Long-Term Evolution (LTE) variants. It is therefore necessary to investigate more advanced network optimizations that improve the last hop wireless connectivity and thereby help the mobile operators to satisfy the need for mobile data.

Initial Problem

In wireless networks the location of a user often has a big influence on the network performance, for example in relation to coverage regions. For example user movements are often the underlying triggers for network reconfigurations. In traditional network optimizations it is often link quality measurements that are used to trigger reconfigurations. Since location information relates directly to the user's location or movements, which are causing the need for reconfiguration, it may in some cases be beneficial to use location information as input for network optimizations. On the other hand, since location information does not accurately describe the actual state of the wireless channel, there may be cases where the location information does not provide any additional benefits. One challenge to address in this work is therefore to investigate under which circumstances location information based network optimizations are useful.

Network optimizations can be classified according to the time scales they are operating on. On the small time scale, optimizations that work on for example individual transmission symbols, are not well suited for exploiting location information, since location information takes time to gather and process. On large time scales that cover hours or days, optimizations such as expanding core network capacity by renting leased lines to meet increased demands is also not interesting since the dynamics of mobile users work on smaller time scales. On the medium scale, which covers fractions of seconds to minutes, there is a potential for location information to be useful for optimizing network functionalities according to the dynamics of mobile users. Furthermore, in order to exploit the advantage of location information over local link quality measurements, which is the knowledge of the geographical relations of users, the

considered network optimizations should focus on distributed systems where this knowledge is exploited.

Seeing this delimitation in the light of the need for improved last hop wireless connectivity, a promising focus area of this work is *last hop route selection*. This targets the problem of determining the best configuration for connecting a mobile user to the network operator's infrastructure. Given the mobile operator's future challenge of meeting the costumers' rising demand for mobile data, this work is considering the *example use case of cellular network offloading* for the study of location based network optimizations.

Cellular Network Off-loading

A solution that allows mobile operators to keep up with the demand would be to extend their networks with Wi-Fi hotspots and allow users to connect to these high-speed hotspots when in range. This would then move some of the mobile data traffic to Wi-Fi network and thereby offload the cellular networks. From the user's point of view, this solution has several advantages. Since cellular data plans are typically charged per MB or allow a certain amount of mobile data each month, connecting to Wi-Fi hotspots when available will allow the user to save on their mobile data plan.

Another important factor, which has an influence on the user experience, is the number of users being served within each cell or by a Wi-Fi Access Point (AP). Typically, the radius of a single cell in a cellular network ranges from several hundreds of meters up to tens of kilometers. In comparison, the coverage range of a Wi-Fi hotspot is typically below 100 meters. Since a cell covers a much larger area than a Wi-Fi hotspot, the number of users sharing the capacity within a cell is therefore much higher. As a result hereof, a Wi-Fi connection will likely provide a higher throughput and lower latency for the user applications than the cellular network.

In summary, off-loading the cellular networks using Wi-Fi networks is advantageous for both the mobile operators and the mobile users as it relieves the stress on the operators' networks and improves the throughput for the users.

The foreseen need for off-loading the cellular networks using Wi-Fi networks and the ubiquitous availability of location information in mobile devices, has lead to the main research question addressed in this thesis:

How well suited are location based network optimizations for wireless last hop route selection in networks with mobile users?

1.2 Scenarios

In the present work this main problem will be approached from two different viewpoints: large and small scale.

On a large scale, as exemplified by the scenario in Figure 1.1, the availability of different access networks motivates the investigation of location-based handover algorithms that can determine when and where to use which access network. The potential of such algorithms is shown in the two experimental performance comparisons of 3G and Wi-Fi in [Gass and Diot, 2010] and [Deshpande et al., 2010]. From these it is clear that if a terminal is able to know which throughput is achievable for each available network at specific locations, then the terminal could use the knowledge of its own location to decide which network to handover to. The mapping of specific locations to achievable network throughput, could be done through a central database, which is populated by the operators, customers' equipment by using applications such as cellumap[1], or by 3rd parties such as Google. Further, assuming that the user's movements can be predicted with high certainty for some time into the future, network handovers could be planned and prepared in advance. The user's expected future movements could be known for example from a navigation system, by learning the user's usual routes, or by taking into account road layouts. Since such a location based handover algorithm relies on location information and movement prediction, the accuracy of the location information and the accuracy of the movement prediction algorithms are determining factors for the performance of the handover scheme.

On the small scale, when mobile users are connected to a single Wi-Fi access point, e.g., for offloading, as exemplified in Figure 1.2, it is well known that relaying techniques, where neighboring devices are used as relays can be used to enhance overall performance as described in [Narayanan and Panwar, 2007]. Relaying is for example helpful in the case where nodes located far away from the access point may be forced to use a low bit rate to achieve a reliable link, as indicated in Figure 1.2, where the darkness of the circles around the access point corresponds to different bit rates. By introducing relaying techniques, nodes located between a transmitter and receiver pair can be exploited to provide a multihop path with shorter links. The shorter links between nodes may form a more reliable path to the destination node, which allows to increase the bit rate and achieve a higher throughput, see for example [Zhu and Cao, 2006, Liu et al., 2007, Hu and Tham, 2010]. By improving the performance within Wi-Fi hot spots, the cellular network can be offloaded more than if

[1] See `www.cellumap.com`.

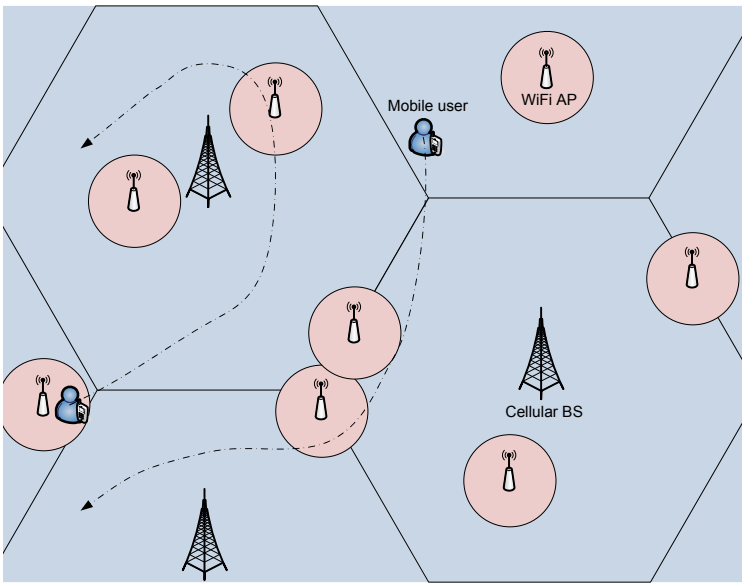


Figure 1.1: Large scale scenario with ubiquitous cellular coverage and scattered Wi-Fi access points. The two mobile users may potentially achieve a higher throughput if they handover to the Wi-Fi networks, when in range. The dash-dotted lines indicate possible movement trajectories.

relaying was not used. As these existing schemes rely on past measurements

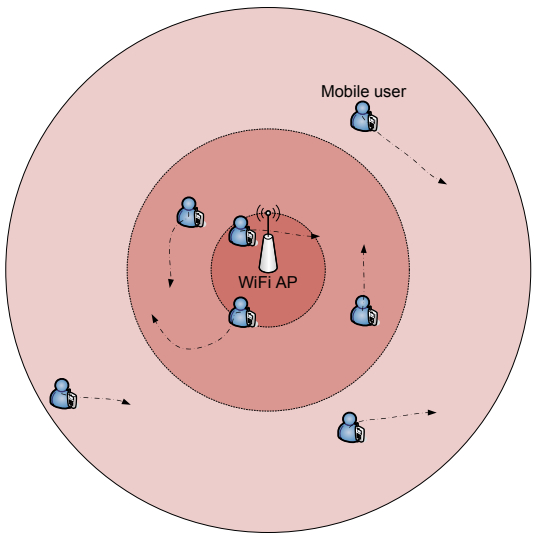


Figure 1.2: Small scale scenario considering only the users being served by a single access point. The colored rings around the Wi-Fi access points represent the coverage area for different bit rates. The dash-dotted lines indicate possible movement trajectories.

of link quality for relay selection, they do not work well in mobile settings where nodes move around and measurements become outdated, as shown in [Liu et al., 2007]. In such mobile settings, it may therefore be beneficial to consider a location based approach for relay selection, as location information can be used to estimate link qualities by accounting for the distance dependent path loss. The actual quality of a link is also depending on other factors such as multi-path fading or obstructions. This is not captured by a location based approach, but it is captured for the measurement based approaches mentioned above. However, if in mobile settings, the measurements become outdated too often, it may be better to use up-to-date but less precise location information instead. In order to determine when a location based scheme is preferable over a measurement based scheme, a study that accounts for factors such as collection delay, information inaccuracy, and node mobility is needed.

In two-hop relaying approaches, a relay-to-destination transmission, which follows the AP-to-relay transmission, will typically take place quite far from the nodes associated with the AP that are located on the opposite side of the AP. This has been exploited in the CCMAC protocol described in [Hu and Tham, 2010], where spatially separated uplink source-to-relay transmissions are performed simultaneously, whereafter the relays finalize the two-hop transmissions by transmitting to the AP one after another. Since this protocol uses past measurements to coordinate the relay selection and to determine which transmissions can occur simultaneously, it uses a learning algorithm that learns the best configuration over time. Again, mobility would be problematic for this protocol, since it takes time for it to identify suitable relays and candidates for simultaneous transmissions. In the same way as location information can be used to estimate link quality via a path loss model, it can also be used to predict interference, and would thus be useful for identifying suitable relays and candidates for simultaneous transmissions. As for the location based relay selection discussed above, it would be necessary to study how such a location based scheme for simultaneous transmissions reacts to inaccurate information, specifically inaccuracy location information.

A general prerequisite for the considered location based protocols, is the provision of location information. As GPS based localization is not generally usable, e.g., indoors or in urban canyons, it is assumed that a network based localization system is responsible for providing location estimates. In order to judge the performance of the considered protocols, a study of the achievable localization accuracy for such a system in the considered Wi-Fi based scenarios is necessary.

1.3 Network Architecture

The network architecture considered in this thesis is shown in Figure 1.3. This architecture includes the entities that are required for localization as well as location based relaying and location based handover.

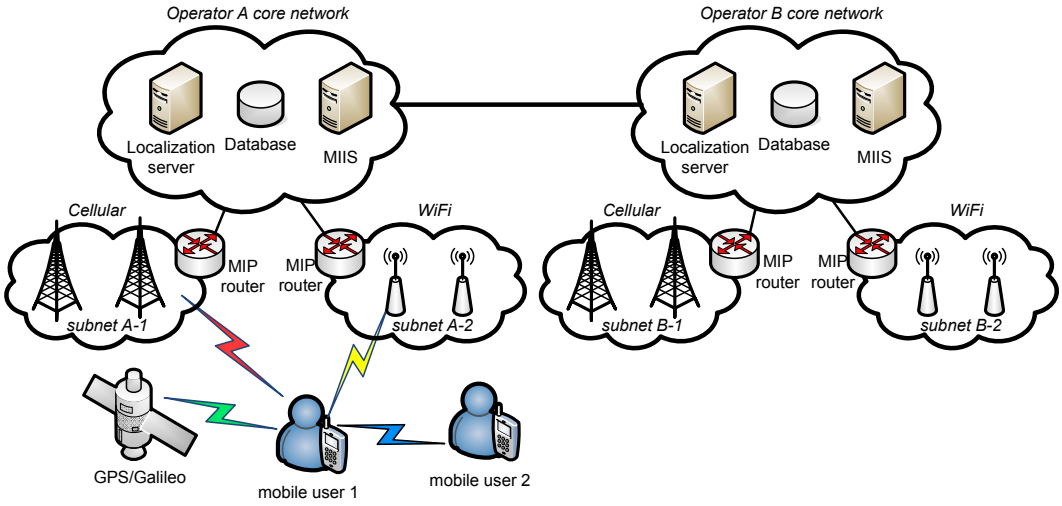


Figure 1.3: Network architecture overview.

First and foremost, the figure shows that a mobile user is able to connect to a mobile operator's core network either via the base stations in the cellular network or via Wi-Fi access points.

For localization, it is assumed that the network operator has a localization server in its core network, which collects measurements that are used for localization. These measurements could taken be from GPS or Galileo satellites, or from communication links to cellular base stations or Wi-Fi access points, or from an ad-hoc link to a cooperative mobile user. All these measurements are collected in the localization server, and from the measurements the localization server estimates the location of the mobile users using a suitable localization algorithm. The resulting location estimates can hereafter be requested by whichever entity that needs them.

For location-based handover it is assumed that the IEEE 802.21 Media Independent Handover (MIH) framework is used. This framework helps to realize handovers between heterogeneous networks [Taniuchi et al., 2009]. The protocol defines tools to exchange information, events, and commands to facilitate initiation and preparation of handovers, and requires only a device driver for each type of media (e.g. Wi-Fi, UMTS, or LTE) that implements

the MIH interface. In practice, the MIH functions are added as a layer 2.5 in the protocol stack, where layer 3 protocols are presented with new MIH interfaces [Neves et al., 2009]. The IEEE 802.21 framework does not specify mechanisms for executing handovers, but several have been proposed in the literature [Neves et al., 2009][Yuliang et al., 2010][Lampropoulos et al., 2010] and it is foreseen that the handover algorithms proposed in this thesis can be used within this framework.

The core of MIH is the MIH Functions (MIHF) that is a logical entity located between the media specific link layers and the MIH user, which represents layer 3 and above. As described in [Lampropoulos et al., 2010] the MIHF can be located in both the mobile device and in a network entity, and supports services for subscription and reception of link-related events through the Media-Independent Event Service (MIES), execution of commands for controlling link states through the Media-Independent Command Service (MICS), and retrieval of information regarding available networks through the MIH Information Server (MIIS). According to [Baek et al., 2008] the MIH framework allows handovers to be triggered both from the mobile device or from the network, depending on how the MIHFs are installed in the network. Regardless of where the handover is triggered, the MIIS is needed in the core network. This entity can be queried by other MIH entities to get information about available networks for a certain geographic area, which allows for sophisticated handover decisions. For the location based handover it is assumed that the MIIS has access to a database containing information about the expected throughput for the available cellular and Wi-Fi networks for any geographic location, as this is needed for the location based handover algorithms. Such a database could be made empirically from field-measurements as in [Deshpande et al., 2010] or synthetically using for example ray tracing techniques.

Further, since a handover to another network will cause the IP address to change, it is necessary to use Mobile IP (MIP) to ensure service continuity for user applications [Perkins, 2002]. MIP adds extra functionality to the IP layer (layer 3) which allows the mobile user to keep its IP address even when visiting foreign networks. Whenever the mobile user is not in the home network (where its IP address belongs) a foreign agent in the visiting network will forward the traffic to and from a home agent in the home network, thereby allowing the home IP address to be used.

The MIP functionality resides both in the user device and in the network routers (as shown in Figure 1.3). The foreign agent and home agent roles are filled by the MIP routers located in each network, depending on which network the mobile device is associated with.

The location based relaying functionality that is considered in this thesis, resides within the Wi-Fi access points, since it only concerns the mobile devices associated with that access point. For obtaining the location of the mobile devices, the access point queries the localization server. When considering the relaying functionality, the direct link between the two mobile users in Figure 1.3 represents a relay link.

1.4 Objectives

Based on the above problem definition and scenario descriptions, the core problem of this work is to:

> *Investigate the benefit of using location information for relaying and handover network optimizations as a means to improve the last hop wireless connectivity of mobile users by exploiting available Wi-Fi networks.*

Notice that the overall objective is not to propose relaying and handover protocols that are better than existing state of the art protocols, but rather to study the benefits and drawbacks of using location information compared to traditional schemes using link quality measurements.

Based on the core problem definition above, the objectives of this thesis can be summarized as the following items:

Location based relaying Study the benefits and drawbacks of location based relaying, compared to measurement based relaying for the small scale scenario with mobile users around a Wi-Fi hotspot. The study should especially focus on studying the impact of:

- node movements
- inaccurate location information
- measurement collection (e.g., delay and overhead)
- propagation model simplification, i.e., using path loss model for the location based scheme

Further, since location information enables simple estimation of interference, the feasibility of an extended relaying scheme that allows simultaneous transmissions should by studied. Also here the impact of inaccurate location information is essential.

Location based handover Investigate methods for exploiting the prediction power of location information in handover optimization, in order to allow offloading of cellular network by exploiting available Wi-Fi capacity in hotspots. Prediction power means that a moving node's future movements can be predicted, with some level of certainty, from its past movements. If the locations and transmission power of access networks are known in advance, the availability of networks can be anticipated and handovers can be planned in advance. The location based schemes should be compared to traditional measurement based schemes. Being dependent on location information and movement prediction, the impact of inaccuracies in these parts should be studied.

Achievable localization accuracy Investigate the achievable accuracy of cooperative network based localization systems, in order to realistically assess the impact on the considered location based relaying and handover schemes. Since cooperative localization requires the cooperating devices to exchange measurements, an important aspect to include is the delay caused by this distribution of measurements.

In summary, the objectives of this thesis can be visualized as in Figure 1.4. Here the four different aspects that are investigated for measurement and location based network optimizations are depicted. A description of each of these aspects, shown by the numbered circles, is given below:

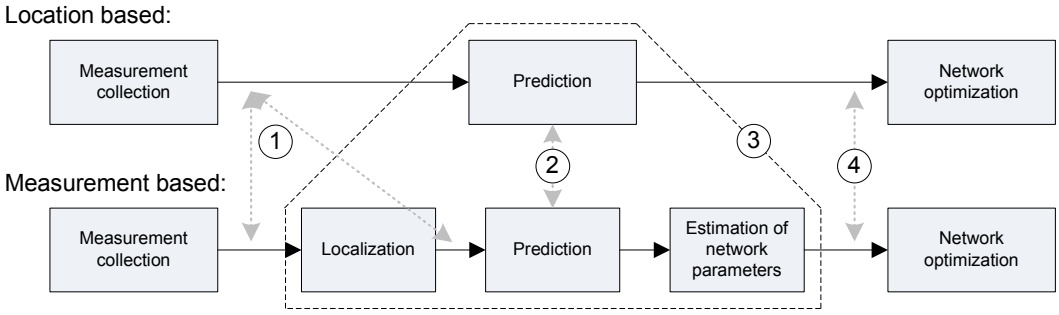


Figure 1.4: Overview of the main objectives of the analyses in the thesis.

1. Signaling overhead for measurement collection may be reduced by using location information. Specifically, only selected link measurements are needed vs. measurements of all links; or only N node locations are required vs. N^2 link measurements.

2. Better prediction of future connectivity options by exploiting a priori information such as mobility models or environment knowledge.

3. Exploitation of geometric relations to estimate outcome of complex situations may be possible for cases where many nodes/links are involved, and where measuring is not possible/feasible due to too many combinations.

4. Quality of input information: The main focus of the thesis is to study under which circumstances the location based approach results in best input information?

1.5 Methodology

In the parts of this work where simulations are used, considerations are made to choose an appropriate level of abstraction as well as the evaluating the required effort of different approaches. Consequently, the simulation results have been generated through a combination of *ns-2* simulations and custom matlab code. *ns-2* has been used in the cases where realistic timing of 802.11 is needed, whereas custom-built matlab simulations were used in the cases where the level of detail in ns-2 was considered too high and a less detailed model would suffice.

In relation to the performed simulations, focus has been on ensuring statistically sound results. This has been achieved by setting the length of simulations and the number of simulation runs so that the 95% confidence intervals are non-overlapping for results that are concluded to be different.

1.6 Contributions

The overall contribution of this PhD work is the investigation of how suitable location based network optimizations are for offloading cellular networks by using Wi-Fi networks. Figure 1.5 gives a detailed overview of the individual contributions and how they relate to the different thesis chapters.

The figure outlines the central components of this work. The considered network optimizations rely on obtaining measurements from the network, which is symbolized by the *measurement collection* block. For the considered location based algorithms, a *localization system* provides location information. These inputs constitute the *observed system state*, from which the considered *optimization algorithms* determine the most suitable *configuration*. In order to evaluate the goodness of different configurations, the optimization algorithms

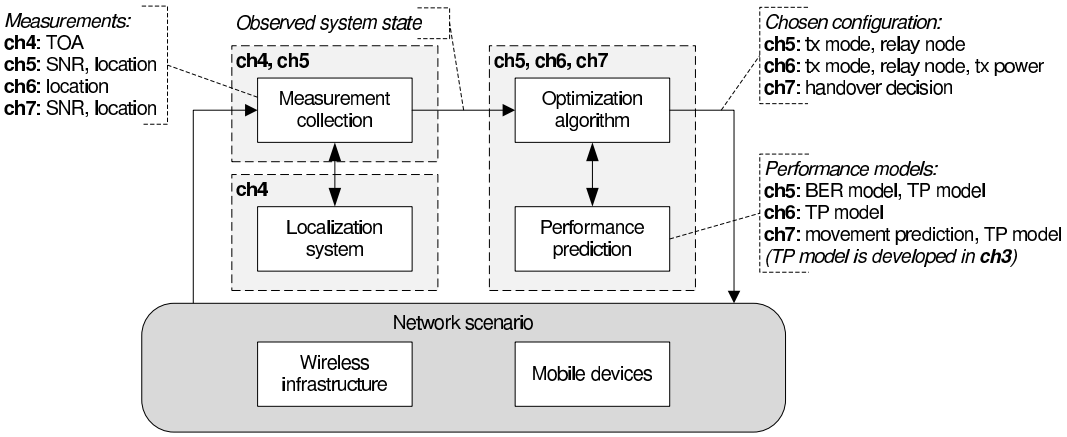


Figure 1.5: Abstract system overview of thesis contributions.

rely on performance models to estimate the achieved system performance for different configurations.

Table 1.1 presents an overview of the types of errors that are considered for the different contribution chapters in the thesis.

Distance-dependent throughput model for Wi-Fi This throughput model (TP model in figure) describes the expected achievable throughput between a transmitter and receiver in a Wi-Fi network, based on the path loss between the nodes. The contribution consists of a formulation of the expected link throughput as a function of received power, interference, bit error rate and frame error rate. The throughput model, which is described in Chapter

Error types	**Ch3**	**Ch4**	**Ch5**	**Ch6**	**Ch7**
node mobility		X	X		X
measurement collection delay		X	X		
small-scale fading	X (BER calc)		X (SNR var, BER calc)	X (BER calc)	X (TP var, BER calc)
inaccurate prop. model par.			X		
interference	X			X	
location accuracy		X (out-come)	X	X	X
movement prediction inaccuracy					X

Table 1.1: Overview of which error types are considered in which chapters.

3, is used as performance model for the location based network optimizations considered in chapters 5, 6, and 7.

Realistic Communication Constraints for Localization This contribution is a study of how realistic communication constraints affect conventional and cooperative localization and tracking algorithms. This contribution, which is described in chapter 4, is a part of a joint work, and serves as a means to understand the achievable location information accuracy of a realistic network based localization system. For the cooperative localization, a group mobility model is used to generate correlated user movements. The personal contributions consist of: i) developing the framework for how to approach the problem; ii) an implementation of an ns-2 simulation model of the message exchanges required to realize conventional and cooperative network-based localization algorithms; and iii) an analysis of the message exchange timing in this simulation model for various scenario parameters.

Location-based Relaying The contribution concerning relaying schemes for centralized selection of mobile relays for two-hop relaying is presented in Chapter 5. Initially, a conventional SNR-measurement based and a location based relaying scheme are defined and compared. This includes implementation of ns-2 and matlab based simulation framework for evaluating the impact of mobility and timing of measurement exchanges on relaying performance. Further, since the location based scheme uses a path loss model to predict performance, the impact of inaccurate path loss model parameters is investigated. In order to deal better with non line of sight situations, a simple extended scheme which has LOS/NLOS information is evaluated.

Simultaneous Transmissions in Relaying This contribution, which is described in Chapter 6, is an interference-aware extension of the location based relaying scheme that increases the downlink throughput by allowing simultaneous relay-to-destination transmissions. The scheme is evaluated with respect to the impact of inaccurate location information on the throughput gain obtained by allowing simultaneous transmissions.

Handover Optimization This last contribution in Chapter 7 is related to location based handover optimization in heterogeneous networks. The problem of deciding when to handover between different available networks within a fixed time horizon has been formulated as an optimization problem. Secondly,

the optimal solution of the handover problem is outlined, assuming continuous differentiability of the functions used to describe the expected throughput. Also a heuristic algorithm, which is feasible for online use, is proposed. The location based algorithms are evaluated numerically and compared to some reference schemes, in order to determine the impact of node mobility, location information inaccuracy, and movement prediction inaccuracy.

These contributions are presented in the following conference articles:

- Jimmy Jessen Nielsen, Tatiana K. Madsen, Hans-Peter Schwefel, *Location Assisted Handover Optimization for Heterogeneous Wireless Networks*, European Wireless conference, Vienna, Austria, 2011
- Jimmy Jessen Nielsen, Tatiana K. Madsen, Hans-Peter Schwefel, *Location-based Relay Selection and Power Adaptation Enabling Simultaneous Transmissions*, IEEE GLOBECOM, MCECN workshop, Miami, Florida, 2010
- Christian Mensing, Jimmy Jessen Nielsen, *Centralized Cooperative Positioning and Tracking with Realistic Communications Constraints*, WPNC, Dresden, Germany, 2010
- Jimmy Jessen Nielsen, Tatiana K. Madsen, Hans-Peter Schwefel, *Location-based Mobile Relay Selection and Impact of Inaccurate Path Loss Model Parameters*, IEEE Wireless Communications and Networking Conference (WCNC), Sydney, Australia, 2010
- Jimmy Jessen Nielsen, Tatiana K. Madsen, Hans-Peter Schwefel, *Mobility Impact on Centralized Selection of Mobile Relays*, IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, Nevada, 2010
- Christian Mensing, Stephan Sand, Jimmy Jessen Nielsen, Benoit Denis, Mickael Maman, Jonathan Rodriguez, Senka Hadzic, Joaquim Bastos, Ziming He, Yi Ma, Santiago Zazo, Vladimir Savic, Igor Arambasic, Mohamed Laaraiedh, Bernard Uguen, *Performance Assessment of Cooperative Positioning Techniques*, ICT Future Networks and Mobile Summit, Florence, Italy, 2010
- Stephan Sand, Jimmy Jessen Nielsen, Christian Mensing, Yi Ma, Rahim Tafazolli, Xuefeng Yin, Joao Figueiras, Bernard H. Fleury, *Hybrid data fusion and cooperative schemes for wireless positioning*, IEEE Vehicular Technology Conference (VTC), 2008

Further, the following articles present work that has contributed indirectly to this thesis on a methodological level.

- Kasper Revsbech, Jimmy Jessen Nielsen, Kevin Harritsø, Rainer Steffen, *Analysis of IP-based Real-Time In-Car Networks with Network Calculus*, The 29th IEEE Real-Time Systems Symposium (RTSS), Barcelona, Spain, 2009
- Jimmy Jessen Nielsen, Lars Jesper Grønbæk, Thibault Renier, Hans-Peter Schwefel, Thomas Toftegaard, *Cross-Layer Optimization of Multipoint Message Broadcast in MANETs*, IEEE Wireless Communications and Networking Conference (WCNC), Budapest, Hungary, 2009.
- Jimmy Jessen Nielsen, Amen Hamdan, Hans-Peter Schwefel, *Markov Chain-based Performance Evaluation of FlexRay Dynamic Segment*, 6th Intl workshop on Real Time Networks, Pisa, Italy, 2007
- Rasmus Løvenstein Olsen, Anders Nickelsen, Jimmy Jessen Nielsen, Hans-Peter Schwefel, Martin Bauer, *Experimental analysis of the influence of context awareness on service discovery in PNs*, 15th IST Mobile & Wireless Communication Summit, Mykonos, Greece, 2006

References

J.Y. Baek, D.J. Kim, Y.J. Suh, E.S. Hwang, and Y.D. Chung. Network-initiated handover based on IEEE 802.21 framework for QoS service continuity in UMTS/802.16 e networks. In *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, pages 2157–2161. IEEE, 2008.

Cisco Systems. Global Mobile Data Traffic Forecast Update, 2009-2014. *Cisco Systems Feb 9th*, 2010.

P. Deshpande, X. Hou, and S.R. Das. Performance Comparison of 3G and Metro-Scale WiFi for Vehicular Network Access. *10th ACM Internet Measurement Conference (IMC 2010), November 1-3*, 2010.

Richard Gass and Christophe Diot. An experimental performance comparison of 3g and wi-fi. In *Passive and Active Measurement*, volume 6032 of *Lecture Notes in Computer Science*, pages 71–80. Springer, 2010.

Zhengqing Hu and Chen-Khong Tham. Ccmac: Coordinated cooperative mac for wireless lans. *Computer Networks*, 54(4):618 – 630, 2010. ISSN 1389-1286. doi: 10.1016/j.comnet.2010.02.001. Advances in Wireless and Mobile Networks.

G. Lampropoulos, C. Skianis, and P. Neves. Optimized fusion of heterogeneous wireless networks based on media-independent handover operations [Accepted from Open Call]. *Wireless Communications, IEEE*, 17(4):78–87, 2010. ISSN 1536-1284.

P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S.S. Panwar. CoopMAC: A cooperative MAC for wireless LANs. *IEEE Journal on Selected Areas in Communications*, 25(2):340, 2007.

S. Narayanan and S.S. Panwar. To Forward or not to Forward–that is the Question. *Wireless Personal Communications*, 43(1):65–87, 2007.

P. Neves, F. Fontes, S. Sargento, M. Melo, and K. Pentikousis. Enhanced media independent handover framework. In *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, pages 1–5. IEEE, 2009.

C.E. Perkins. Mobile ip. *Communications Magazine, IEEE*, 35(5):84–99, 2002. ISSN 0163-6804.

K. Taniuchi, Y. Ohba, V. Fajardo, S. Das, M. Tauil, Y.H. Cheng, A. Dutta, D. Baker, M. Yajnik, and D. Famolari. IEEE 802.21: Media independent

handover: Features, applicability, and realization. *Communications Magazine, IEEE*, 47(1):112–120, 2009. ISSN 0163-6804.

T. Yuliang, L. Chun-Cheng, K. Guannan, and D. Der-Jiunn. Cross-Layer Handover Scheme for Multimedia Communications in Next Generation Wireless Networks. *EURASIP Journal on Wireless Communications and Networking*, 2010, 2010. ISSN 1687-1472.

H. Zhu and G. Cao. rDCF: A relay-enabled medium access control protocol for wireless ad hoc networks. *IEEE Transactions on Mobile Computing*, pages 1201–1214, 2006. ISSN 1536-1233.

2

Background

This chapter gives a brief introduction to some basic concepts that the remaining parts of the thesis are based on. Initially, an overview of the factors that impact the performance of IEEE 802.11 based Wi-Fi systems is given, as it plays a central part in both the small and large scale scenarios. Secondly, the basic concept of relaying is introduced, as it is in focus for the small scale scenario investigations. Next, network based localization is introduced and related to satellite based localization systems. Finally, an overview of the mobility models that are used in this thesis is given.
Notice that the individual contribution chapters contain their own specific state of the art sections.

2.1 Wi-Fi Performance Aspects

The Wi-Fi networks considered in this work are assumed to be based on the IEEE 802.11a protocol, which is described in [IEEE, 2007]. This section briefly introduces the different factors and models that determine the performance of such a Wi-Fi network system. These are used in the later contribution chapters of this thesis.

2.1.1 Path loss

In wireless networks, where the receiving node of a transmission is spatially separated from the transmitting node, as illustrated in Figure 2.1, it is well known that the energy level at the receiver depends on the transmission distance.

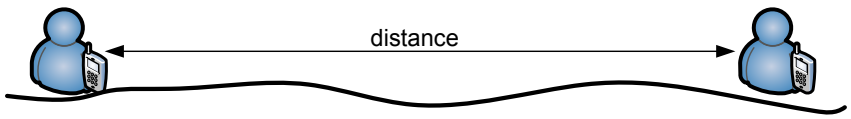


Figure 2.1: Example of path loss over distance.

This phenomenon is called the path loss and can be modeled using a path loss model and knowledge about the transmit power, environment characteristics, and the distance between transmitter and receiver. Some typical path loss models are the following (see for example [Goldsmith, 2005]):

Free space (Friis) model: As the name implies, this model is built on an assumption of free space between the transmitter and receiver. The model is useful for situations where the transmitter or the receiver (or both) are positioned above the ground, and there are no obstacles between them.

Two ray (ground reflection) model: This model is an extension of the free space model in which it is assumed that the received signal consists of a direct (Line-of-Sight (LOS)) radio wave and a single other wave which is reflected on the ground. This model is useful in for example vehicular scenarios, where the road is the reflector.

Log distance (path loss exponent) model: This model calculates the distance dependent path loss given a path loss exponent parameter. The path loss exponent depends on the environment and typically has a value between 2 and 4.

As this work focuses on semi-urban and urban scenarios where sometimes obstructions hinder the direct radio transmission path, neither the free space model or the two ray model are well suited for this work. The log distance model is more versatile since it allows different path loss exponents to be used for different scenarios, and will therefore be used in this work.

The 802.11a Wi-Fi networks considered in this work use the 5 GHz U-NII band. In [Durgin et al., 1998] the authors have made measurements in and around homes and trees for this specific frequency band and determined appropriate path loss exponent values for this type of environment. Since Wi-Fi access points are typically located inside homes or similar buildings, the path loss model applied in this work is based on their identified parameters. The model structure of the log-distance model is:

$$P_{\mathrm{rx}}(d)\ [dBm] = P_{\mathrm{tx}} - PL(d_0) - 10n \log_{10}(d/d_0) \tag{2.1}$$

where P_{tx} is the transmit power, $P_{\mathrm{rx}}(d)$ is the received power at the receiver, d is the distance between transmitter and receiver, $PL(d_0)$ is the path loss at a reference distance $d_0 = 1$m, and n is the path loss exponent. In [Durgin et al., 1998] the authors report values ranging from 2.7 to 3.6 for indoor and outdoor environments around homes.

Figure 2.2 shows an example of the received power relative to a transmit power $P_{\mathrm{tx}} = 20$ dBm, which is the typical transmit power of a Wi-Fi access point, as a function of distance.

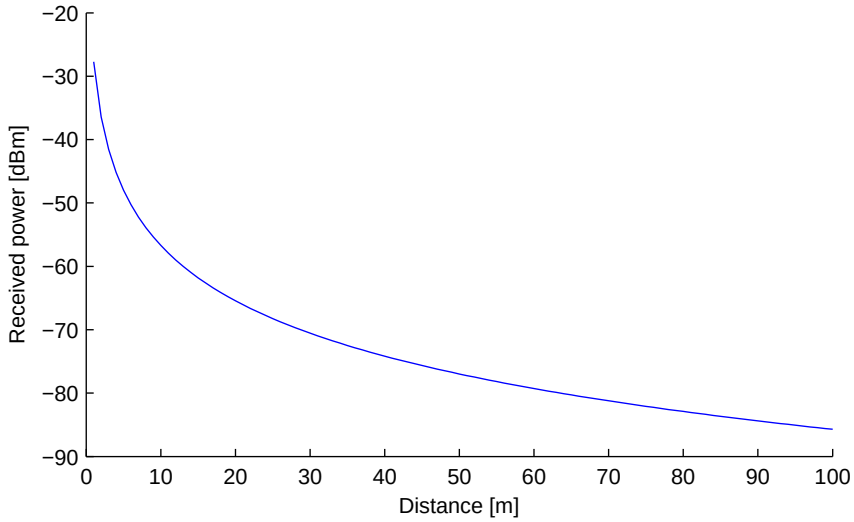


Figure 2.2: Example of received power for $P_{\mathrm{tx}} = 20$ dBm, $PL(d_0) = 0$, and $n = 2.9$.

2.1.2 Multi Path Fading

In scenarios with objects and obstructions such as buildings, walls, vehicles, people, or furniture, radio waves do not only travel along the LOS path between a transmitter and a receiver. The receiver will likely receive a combination of the direct and reflected waves that may partially cancel out or boost the received power at the receiver. This phenomena causes quickly changing variations in the received signal power if the user or other objects are moving. These variations in received power are superimposed on the path loss.

For a digital transmission, the multi path fading has a significant impact on the Bit Error Rate (BER) when deep fades occur. Here, the received power drops to a very low level and the amount of bit errors can be significant. This means that in average a higher SNR is therefore required for a given BER compared to the ideal case of an AWGN channel.

In cases where the received power at the receiver comes mainly from scattered waves, such as indoor or dense urban environments, the received signal can be described using the Rayleigh fading model, see, e.g., [Rappaport, 2002]. This model assumes that the received signal consist of a large number of scattered waves with independent and identically distributed (i.i.d.) inphase and quadrature amplitudes.

In cases where the received signal consists of some scattered waves but with one dominant component, e.g., the LOS component, the Ricean (or Rician) fading model is used to describe the signal, see, e.g., [Rappaport, 2002]. This model is similar to the Rayleigh model except for the dominant component. The model has a parameter, typically denoted as K, which specifies the ratio between the power of the dominant component and local mean of the scattered components. For $K = 0$ ($-\infty$ dB), the Rician model corresponds to the Rayleigh model.

The function `berfading` from the matlab communications toolbox can be used to calculate the achieved BER for Rician or Rayleigh fading environment for different modulation schemes and values of the Rician factor K.

As this work focuses on semi-urban and urban scenarios where only sometimes obstructions hinder the direct radio transmission path, the Ricean fading model is judged to be more appropriate than the Rayleigh fading model.

2.1.3 IEEE 802.11 based Wi-Fi

The IEEE 802.11 standard [IEEE, 2007] specifies the functionality of the physical (PHY) and Medium Access Control (MAC) layers. There are several variants available.

802.11 Physical Layer Variants

802.11b and 802.11g both operate in the ISM band around 2.4 GHz and they support bit rates of up to 11 Mbit/s and 54 Mbit/s, respectively. 802.11a is in many ways similar to 802.11g as it supports bit rates of up to 54 Mbit/s, but it is operating in the 5 GHz U-NII frequency. Finally, the most recent addition is 802.11n, which can operate in both the 2.4 GHz and 5 GHz frequency bands, but uses MIMO technology to increase range and throughput. The standard supports bit rates of up to 600 Mbit/s for certain hardware configurations.

802.11 Medium Access Control

IEEE 802.11 specifies two medium access control protocols, Point Coordination Function (PCF) and Distributed Coordination Function (DCF). PCF is a centralized scheme, whereas DCF is a fully distributed scheme. DCF is the most commonly used scheme. The DCF scheme in IEEE 802.11 is based on a slotted Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol, as is described briefly in the following. For a more detailed description see reference [IEEE, 2007].

DCF is a contention based medium access protocol. When a station has a frame to transmit, it transmits the frame immediately if the channel is idle for the duration of a Distributed InterFrame Space (DIFS). If the channel is busy the station waits until the channel remains idle for the duration of a DIFS. When this happens, the station randomly instantiates its backoff counter on the interval $0, W - 1$, where W is the contention window, which depends on the number of failed transmission attempts. For each time slot in which the channel remains idle the station decrements its backoff counter. If the channel becomes busy, the station freezes its backoff counter and waits until the channel has been idle for a DIFS time, before it resumes decrementing the counter. When the counter reaches zero, the station starts its transmission in the beginning of the following timeslot. Here it may happen that another station is transmitting in the exact same time slot, and a collision occurs.

DCF supports two different modes for channel access, the basic mode, in which the receiving station acknowledges every successful reception of a data frame by sending an ACK frame after the duration of a Short InterFrame Space (SIFS), and a four-way handshake mechanism denoted Request to Send / Clear to Send (RTS/CTS), in which the source station reserves the channel before transmitting a data frame. The procedure for the RTS/CTS scheme is as follows. First the source station transmits an RTS frame and after the duration of a SIFS and in the case of no collisions, the receiving station sends out a CTS frame. At this point, all stations that are within hearing range of either the source or the receiving stations have overheard the RTS or CTS frame transmissions and will know that a transmission is ongoing. Since the RTS and CTS frames contain the duration of the data transmission, other stations even know for how long the channel is busy. After another SIFS time, the source station starts transmitting the data frame. When the source station has finished the transmission, the receiving station waits for the duration of a SIFS, and then sends an ACK frame if the data frame was successfully received. Compared to the basic scheme, the RTS/CTS scheme has the advantage that the frames involved in the contention process are much shorter than for the basic scheme, which results in better performance when large data frames are used.

During the process described above, errors can occur either from collisions, where two or more stations within hearing range attempt to use the same time slot, or if the signal quality is not sufficient for the receiver to successfully decode the received signal. In either case, whenever a transmission error occurs, the source station doubles the contention window W, up to a maximum value of $CW_{\max} = 2^m \, CW_{\min}$. The values of $CW_{\min}$ and $CW_{\max}$ depend on the physical layer and can be found in [IEEE, 2007]. For 802.11a, they are $CW_{\min} = 15$ and $CW_{\max} = 1023$.

2.2 Relaying

The basic problem that relaying protocols are helping to mitigate is low signal quality due to fading, including both large scale path loss and small scale multi path fading. A basic three node relay system is shown in Figure 2.3. The idea of a relaying is that a transmission from the source node to the destination node, can also be transmitted from the source S via the relay R to the destination D, which may lead to better reception at the destination. In relation to the path loss, it is clear that each of the links S-R and R-D are shorter than the link S-D and therefore are likely to experience less path

loss. In cases where the link S-D has very low quality due to path loss, it can therefore be beneficial to transmit via the relay R. Further, there may be cases where multi path propagation leads to deep fades on the S-D link even if the path loss in itself is not critical. As the relay R is spatially separated from the destination D, it experiences multi path propagation independently from D, and the relayed path via R may therefore have better link quality.

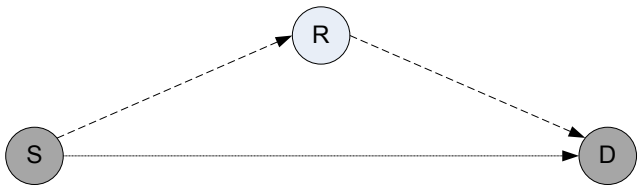


Figure 2.3: Basic three node relay system consisting of source (S), relay (R) and destination (D).

The relayed transmission can be realized in different ways. Some fundamental relaying schemes are (see for example [Nosratinia et al., 2004, Kramer et al., 2006]):

Amplify-Forward (AF) is a scheme in which the relay receives a noisy version of the signal transmitted by the source, which it then retransmits. Even though the relayed transmission contains also amplified noise, the destination has two independently faded copies of the same transmission. One issue with this scheme however is that sampling, amplifying, and retransmitting an analog signal with a sufficient resolution is practically challenging.

Decode-Forward (DF) assumes that the relay attempts to decode the transmitted signal before re-encoding and retransmitting. Compared to AF, this scheme has the advantage that the signal that is retransmitted by the relay does not contain noise. Since the destination may overhear both the source and relay transmissions, it is able to exploit the cooperative diversity for error correction.

Compress-Forward (CF) works by having the relay sample and compress the received signal, whereafter this compressed signal is forwarded to the destination. Like the AF scheme, the sampling and compression of a signal is practically challenging.

Classic multi-hop is similar to DF, however in classic multi-hop the destination does not exploit cooperative diversity and it only decodes the relay transmission.

Since the cooperative AF, DF, and CF schemes cannot be used with state-of-the art wireless network adaptors without modifying the firmware (as described in [Valentin et al., 2009], the scheme that is considered in this thesis is the classic multi-hop.

In Figure 2.3 the considered system consists of the source, destination, and a single relay node. In practical scenarios there are however often multiple potential relay nodes. A big challenge is therefore to determine the most suited relay node in a given situation. This challenge is considered in detail in Chapter 5, where also an overview of existing solutions is given.

2.3 Network-based Localization

Initially motivated by the Federal Communications Commission (FCC) requirement of locating 112/911 emergency calls from mobile phones within tens of meters as described in [Sun et al., 2005], localization of mobile users has attracted increasing attention within the research community. Furthermore, many applications have emerged or are foreseen that exploit position information. Examples include car navigation, intelligent traffic systems, location based advertisement, and indoor guidance systems in airports and warehouses. Such applications have requirements to the position accuracy in the order of tens of meters down to a few meters for the indoor systems. Also timing is an important issue in such applications and requirements in terms of timely accuracy of a location estimate is in the order of several seconds down to below a second for car navigation.

In the case with car navigation, currently available Global Positioning System (GPS) receivers are able to fulfill the localization accuracy requirements whenever sufficient satellites are visible to the GPS receiver. However, in certain cases where the view to the sky is limited, such as in an urban canyon where tall buildings block the GPS signal propagation, sufficient accuracy is not always achieved.

Figure 2.4 shows the Cumulative Distribution Function (cdf) of the number of visible satellites from the Global Navigation Satellite Systems (GNSSs) GPS and Galileo. The CDF is based on simulations of an urban canyon from [Mensing et al., 2008] and exemplifies that the required amount of 4 visible satellites is seldomly achieved even when considering GPS and Galileo systems jointly. An example of the typical localization performance in such an urban scenario is shown in Figure 2.5. Here, it is shown how cellular networks can be used to improve the localization accuracy by fusing distance measurements

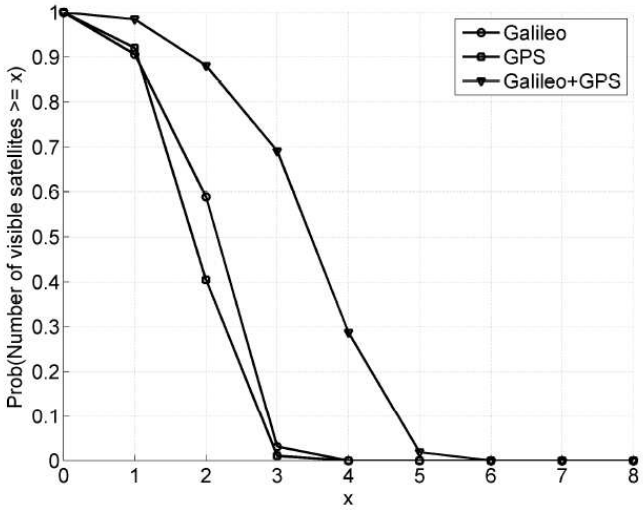


Figure 2.4: CDF of number of visible satellites (x) in urban canyon from [Mensing et al., 2008].

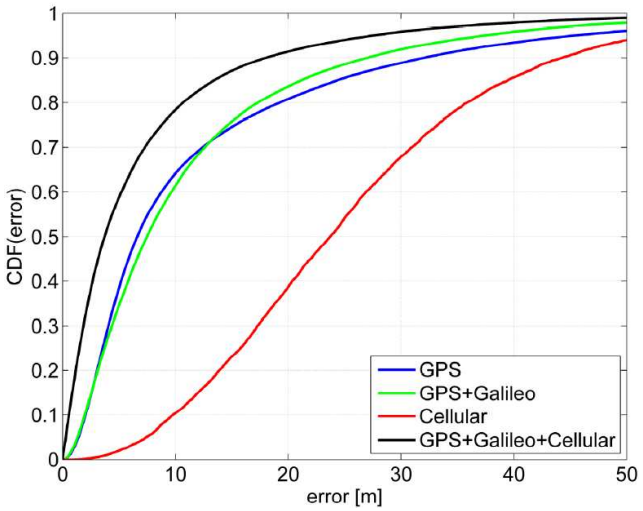


Figure 2.5: CDF of the positioning error in urban canyon from [Pedersen et al., 2010].

to cell towers with GNSS estimates. Furthermore, GNSS based localization rarely works indoor. Therefore, other solutions are needed for location based applications in indoor and dense urban environments.

Recently, the increasing availability of wireless networks has lead to a growing interest for localization solutions based on wireless networks. An example is the Apple iPhone, which uses a hybrid localization solution from the US

company Skyhook[1], in which measurements from both GPS and WiFi access points are used to calculate the position of the user. This approach exploits the presence of the ubiquitous wireless networks that anyone who has attempted to connect his laptop wirelessly in an urban area has most likely noticed, to enable localization where satellite based localization does not work, i.e. indoor and in urban canyons.

Notice that in the literature and also in this thesis, the terms *localization* or localisation and *positioning* are used interchangeably to describe the action of determining the location of an object. In the following subsections different methods for localization in wireless networks are introduced.

2.3.1 Geometric Localization Methods

Typical approaches for wireless localization are based on simple geometrical principles. Different types of measurements may be available, depending on the type of wireless network, and therefore different methods are applied. Some commonly used localization methods are (see, e.g., [Gustafsson and Gunnarsson, 2005, Sayed et al., 2005]):

Trilateration uses ranging measurements such as Received Signal Strength (RSS) or Time of Arrival (TOA). These types of measurements allows the distance between entities, e.g., between the MD and BS, to be roughly estimated. The distance can be estimated from RSS measurements by using a path loss model, if the transmitted signal strength as well as the path loss exponent is known at the receiver. Since the presence of walls and obstructions affects the attenuation of radio signals, the receiver also needs to know or make assumptions of the properties of the surroundings for accurate results in such environments. TOA measurements allow the MD to measure the propagation time of the transmitted signal, which can be used to estimate the distance between a BS and the MD. This approach however requires the clock of the BSs and MD to be very precisely synchronized, which may be a limiting factor in some cases. Based on the estimated distances from ranging measurements, the estimated location of for example an MD can be determined by using trilateration. The estimated location is given as the most likely crossing point of spheres, with radii corresponding to estimated ranges. This is typically solved as an optimization problem since the inaccuracy of measurements means that there is rarely a single unique solution.

[1] www.skyhookwireless.com

Multilateration uses range difference measurements such as Time Difference of Arrival (TDOA) measurements. In practice these measurements are typically easier obtain than TOA measurements, since this approach only requires clock synchronization between BSs, which are in many cases interconnected through a wired infrastructure network. The estimated location of an MD is determined by using the multilateration technique, where the estimated location is given as the most likely crossing points of hyperbolas defined from the TDOA measurements. Also here the solution is typically determined using optimization methods.

Triangulation is as the name implies, based on the use of angle measurements such as Angle of Arrival (AOA). These measurements are obtained by using directional antennas to estimate the angle from which a signal arrives. The estimated location of an MD is determined from AOA measurements using the triangulation method. Like for the two other measurement types, the solution is determined using optimization methods.

In non-ideal cases such as indoor or urban environments, which cover the majority of realistic applications, the estimated position of the MD is not unambiguously given from the measurements due to obstruction or reflection of the radio signals. Therefore, localization approaches usually include error minimization algorithms based on for example least squares or maximum likelihood approaches, or state estimation filters such as the Kalman filter, as described by [Gustafsson and Gunnarsson, 2005]. A supplementing means for decreasing the localization error is to obtain additional measurements from other BSs where the radio propagation path is less disturbed, but this is not always possible.

2.3.2 Fingerprinting

Localization by fingerprinting is a technique which differs from the above by not being based on geometric principles. Instead the fingerprinting technique exploits that different locations typically have a unique fingerprint with regard to, e.g., the visible APs and their respective RSS. In order to work, the fingerprinting technique requires that a fingerprinting database is established that contains an adequate set of fingerprints for the considered area. How many measurements are necessary depend on many factors such as the required accuracy, the BS deployment density, the environment properties, and the measurement grid size. An advantage of this approach compared to the

geometrical approaches listed above is that it can adapt to challenging environments such as indoor environments where walls and building geometry may lead to erroneous ranging or angle measurements. On the other hand this adaptivity also means that if the environment changes in ways that affect the fingerprint, a recalibration of the system is necessary.

2.3.3 Cooperative Localization

Commonly for the approaches mentioned above is that they only consider *local* measurements, i.e. measurements between the BSs and the MD being localized. An example is shown in Figure 2.6 (A) where a Mobile Device (MD) obtains measurements from the three Base Stations (BS) with known locations to estimate its own location.

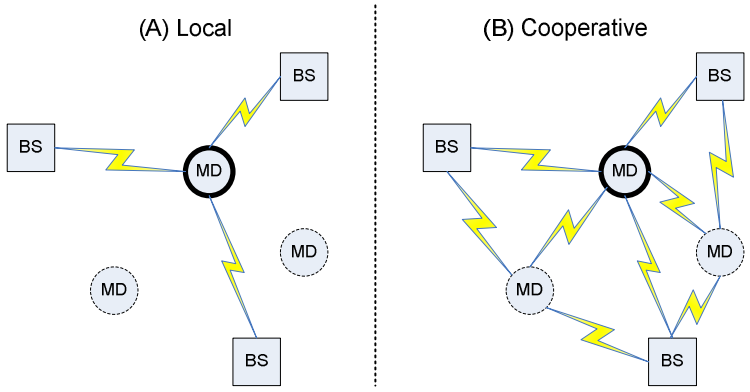


Figure 2.6: Principle of basic local (A) and advanced cooperative (B) localization.

However, since the radio propagation characteristics and thereby localization accuracy is affected by the walls and obstructions that are natural parts of indoor and urban environments it is important to exploit as many measurements as possible, as argued in [Sayed et al., 2005]. This has motivated the step from conventional approaches, where only the links between anchor nodes and mobile nodes are considered for localization, to cooperative approaches where also the links between mobile nodes are exploited for localization. This principle is illustrated in Figure 2.6 (B). In this example, the considered MD, which has a bold outline in the figure, estimates its distance to 5 other nodes and not only the 3 BSs as in part (A) of the figure. The true locations of the 2 MDs are however not known, so their estimated positions may be inaccurate to some degree. But, as shown in the results in Figure 2.7 from [Figueiras, 2008], the added spatial diversity that additional measurements provides, helps to

improve the accuracy of the localization algorithm. This is despite the inaccuracy resulting from the relative MD position estimates.

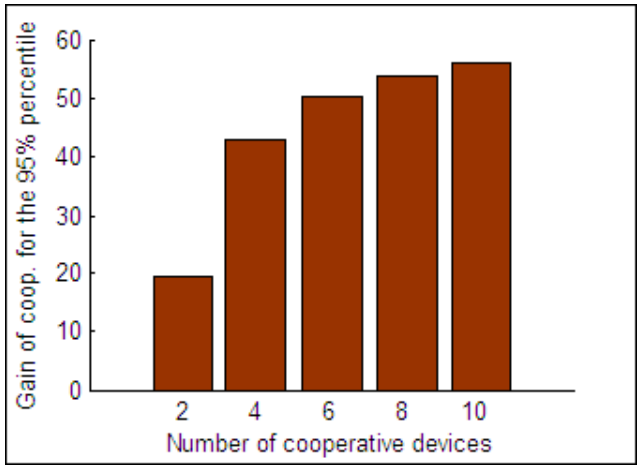


Figure 2.7: Accuracy benefit of additional cooperating nodes, from [Figueiras, 2008].

Basically, cooperative localization can be realized in two ways:

1. Using a centralized approach as in [Mayorga et al., 2007, Frattasi, 2007], where all measurements are collected in one central localization server, which is responsible for jointly processing the measurements to compute the needed location estimates. The location estimates are typically calculated using the geometric methods, depending on the type of measurements. The location estimates can then be exploited in the network (e.g., for optimization of communication functions) or be sent back to the mobile devices. This approach is well suited for infrastructure based systems, where the central localization server can always be reached.

2. The other approach is to use distributed algorithms based on for example Bayesian inference [Wymeersch et al., 2009], where the individual mobile devices are responsible for computing their own location, based on information they share only with their surrounding neighbors via Peer-to-Peer (P2P) links, as in [Wymeersch et al., 2009, Chan and So, 2009]. Since this type of algorithms does not rely on a central entity for processing, these algorithms are well suited for ad hoc networks, such as for example wireless sensor networks.

As this thesis focuses on infrastructure oriented Wi-Fi and cellular networks with mobile users, the most natural choice of localization algorithm would be a

centralized algorithm, since it has all measurements available for processing in the localization server and therefore can achieve higher localization accuracy. Further, the geometric methods are preferred over fingerprinting methods in semi-urban and urban scenarios where the environment cannot be assumed to be static. Chapter 4 presents a study of the achievable accuracy of two centralized geometric localization schemes, where one is based on only local measurements and the other is a cooperative scheme that exploits the measurements from P2P links between users.

2.4 Mobility Models

The purpose of a mobility model is to describe the movements of real world users, in order to study the impact of user movements on system performance. In this work mobility models are used in for simulations to imitate realistic user behavior and study the impact of mobility on localization accuracy, relay node selection, and handover network selection. This section first provides an overview of different types of mobility models, and hereafter the models that have been chosen for this work are described.

In the literature there are various surveys on mobility models such as [Bai and Helmy, 2004, Bettstetter, 2001, Camp et al., 2002, Borrel et al., 2006, Harri et al., 2009, Musolesi and Mascolo, 2009]. In the following, we use the categorization of mobility models presented in [Bai and Helmy, 2004]:

Random Models is the group of mobility models that are used to generate movement patterns, in which the movement parameters are selected randomly, with no dependence on previous selections. This category covers for example the Random Waypoint model, the Random Direction model, and the Random Walk model.

Models with Temporal Dependencies are the models that take into account for example the physical limitations of acceleration, velocity and directionality changes that restrict real world user movements. Such models are especially useful for generating vehicular behavior. Examples of models that fit within this category are the Gauss-Markov mobility model, the Smooth Random mobility model, and the Wrap-around mobility model (see [Haas, 1997]).

Models with Spatial Dependencies are the models in which the user movements are spatially correlated. Such mobility models are useful for generating group-like behavior. Some examples of group mobility models

are the Reference Point Group mobility model, Column mobility model, Pursue mobility model, and the Nomadic Community mobility model.

Models with Geographic Restrictions are used for generating user movements that obey restrictions such as street or building layouts. Examples of mobility models are the Freeway, Manhattan and Pathway mobility models described in [Bai and Helmy, 2004] as well as the spider and Voronoi mobility models described in [Harri et al., 2009].

For the small scale scenario considered in this thesis, which is used for the relaying investigations the Random Waypoint mobility model is used. The primary reason for choosing this model is that related work such as the CoopMAC protocol described in [Liu et al., 2007] presents results where the Random Waypoint model has been used. In order to allow comparison, the relaying investigations in Chapter 5 are based on this mobility model. This model is described in details in section 2.4.1.

For the investigations of cooperative localization in Chapter 4, a group mobility model in which user movements are spatially correlated is needed to imitate groups of users that are willing to cooperate. For this purpose it has been decided to use a Reference Point Group mobility model, which is simply an extension of the Random Waypoint mobility model for the multi-user case. This model is presented in section 2.4.2.

One drawback of the RWP model is that user movements are bounded by the edges of the considered environment, which leads to unnatural behavior of the users in these regions [Borrel et al., 2006]. A way to overcome this, is to use a boundless simulation area/wrap around mobility model such as the one described in [Haas, 1997]. Further this model also has temporal dependencies on the allowed acceleration and direction changes, which leads to a higher level of physical realism. Since the considered handover optimizations in Chapter 7 rely on movement prediction, it is relevant to consider a mobility model that generates realistic movement patterns. This model is therefore used for these handover investigations. The model is described in details in section 2.4.3.

2.4.1 Random Waypoint Mobility Model

This mobility model generates user movements between different waypoints. Given an initial position of the user, the following steps are repeated to produce user movements [Camp et al., 2002]:

- Pause and stay in the current position for a random time p, where p is uniformly chosen in $p_{\min} \leq p \leq p_{\max}$.
- Pick the next waypoint (x, y) randomly (uniform) within the considered area.
- Select a traveling speed v randomly (uniform) in the interval $v_{\min} \leq v \leq v_{\max}$.
- Move in a straight line to the next waypoint at the selected speed v.

The model can be simplified by having no pauses (i.e., $p = 0$). The result is a continuously moving user that changes direction and speed at each waypoint. Another simplification is to set a constant speed, (i.e., $v_{\min} = v_{\max}$). By not using pauses and using a constant speed v, the average speed will always be equal to v.

Figure 2.8 shows an example of a simulation movement trajectory.

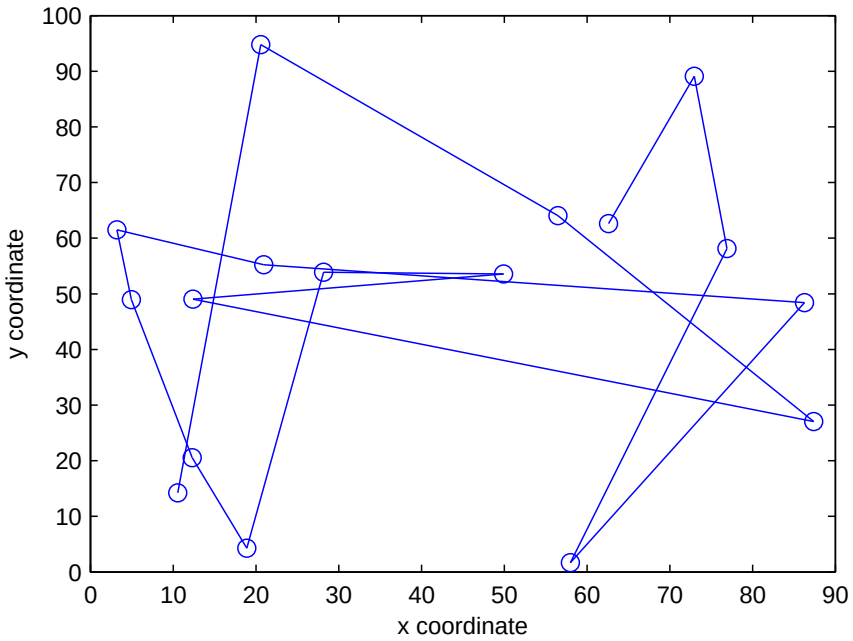


Figure 2.8: Example RWP mobility model outcome for a constant speed of $2m/s$ and $300s$ simulation time. The circles show the waypoints.

In [Yoon et al.] the authors have analyzed the RWP model, and shown that the minimum speed parameter $v_{\min}$ must be set carefully. First and foremost, it is important to select a non-zero value of $v_{\min}$. If $v_{\min} = 0$, the average speed in a simulation will decay and settle at zero. Further, the authors argue that

$v_{\min}$ should not be set too close to zero, as this leads to very long convergence times in simulations.

Another point worth noticing is that the stationary node distribution of the RWP model is non-uniform [Camp et al., 2002]. This means that if the initial user distribution in a simulation study is uniform, there will be an initialization/settling phase before the stationary distribution is achieved. Long simulation runs are therefore needed before this initialization effect is no longer significantly reflected in the results. Alternatively, the initialization phase can be discarded from the results.

2.4.2 Random Waypoint Group Mobility Model

Considering the RWP mobility model, it is fairly easy to introduce correlated movements of multiple users by using the reference point group mobility principle, where one user is defined as a reference and the remaining users are followers, moving in relation to the reference user. Specifically, by selecting the next waypoint for the followers randomly around the reference user's next waypoint, group movements are achieved. That is, each following node's next waypoint should be selected as:

$$\begin{aligned} x_n(i+1) &= x_1(i+1) + \texttt{unif}\left(-\frac{s_{\max}}{2}, \frac{s_{\max}}{2}\right) \\ y_n(i+1) &= y_1(i+1) + \texttt{unif}\left(-\frac{s_{\max}}{2}, \frac{s_{\max}}{2}\right) \end{aligned} \tag{2.2}$$

where $x_n(i+1)$ denotes the waypoint for node n at time index $i+1$, $s_{\max}$ is the maximum spread around the reference node.

An example of the outcome of such mobility model is shown in Figure 2.9, where six groups, each with one reference node and three followers are shown. Here the followers' waypoints are chosen randomly around the respective reference user's waypoint, according to eq. (2.2) with $s_{\max} = 10$.

2.4.3 Wrap-around Mobility Model

This mobility model is based on the model presented in [Haas, 1997], which is referenced in the mobility model survey in [Camp et al., 2002]. This model wraps around, meaning that there are no borders, as shown in [Camp et al., 2002]. This model is used for the large scale scenario in order to avoid issues with coverage on the boundary of the scenario. The implemented mobility

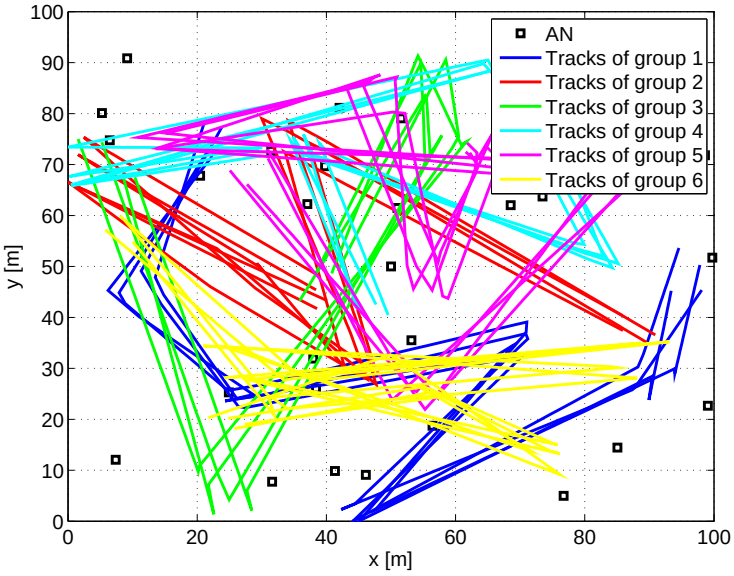


Figure 2.9: Group mobility simulation example, which is used in chapter 4.

model corresponds to [Haas, 1997], however we use the following slightly different formula for calculating the shortest distance between two entities located at (x_1, y_1) and (x_2, y_2):

$$D_x = \min[|x_1 - x_2|, \min(|x_\mathrm{m} - x_1 + x_2|, |x_\mathrm{m} + x_1 - x_2|)]$$
$$D_y = \min[|y_1 - y_2|, \min(|y_\mathrm{m} - y_1 + y_2|, |y_\mathrm{m} + y_1 - y_2|)]$$
$$D = \sqrt{{D_x}^2 + {D_y}^2} \tag{2.3}$$

where x_m and y_m are the horizontal and vertical lengths of the considered area, respectively.

A mobile node's movement is described by its velocity vector $V = (v, \theta)$, where v is the speed and θ is the direction. The location $X = (x, y)$ and velocity V are updated every Δt time as follows:

$$v(t + \Delta t) = \min[\max(v(t) + \Delta v, v_\mathrm{min}), v_\mathrm{max}], \tag{2.4}$$
$$\theta(t + \Delta t) = \theta(t) + \Delta\theta, \tag{2.5}$$
$$X(t + \Delta t) = X(t) + V(t) \cdot \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix}, \tag{2.6}$$

where v_min and v_max are the mobile node's minimum and maximum speeds. The velocity change Δv is a uniformly distributed random variable within

$[-a_{\max} \cdot \Delta t, a_{\max} \cdot \Delta t]$, where $a_{\max}$ is the maximum acceleration/deceleration. The change in direction, $\Delta\theta$ is uniformly distributed within $[-\alpha_{\max} \cdot \Delta t, \alpha_{\max} \cdot \Delta t]$, where $\alpha_{\max}$ is the maximum change in angular speed.

References

F. Bai and A. Helmy. A survey of mobility models in wireless adhoc networks. 2004.

C. Bettstetter. Mobility modeling in wireless networks: categorization, smooth movement, and border effects. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(3):55–66, 2001. ISSN 1559-1662.

V. Borrel, M. de Amorim, and S. Fdida. On natural mobility models. *Autonomic Communication*, pages 243–253, 2006.

T. Camp, J. Boleng, and V. Davies. A survey of mobility models for ad hoc network research. *Wireless communications and mobile computing*, 2(5): 483–502, 2002.

F. W. C. Chan and H. C. So. Accurate distributed range-based positioning algorithm for wireless sensor networks. *IEEE Transactions on Signal Processing*, 57(10):4100–4105, October 2009.

G. Durgin, TS Rappaport, and H. Xu. Measurements and models for radio path loss and penetration loss inand around homes and trees at 5.85 GHz. *IEEE Transactions on Communications*, 46(11):1484–1496, 1998.

Joao Carlos Prazeres Figueiras. *Accuracy Enhancements for Positioning of Mobile Devices in Wireless Communication Networks*. PhD thesis, Aalborg University, 2008.

S. Frattasi. *Link Layer Techniques Enabling Cooperation in Fourth Generation Wireless Networks*. PhD thesis, Aalborg University, Aalborg, Denmark, September 2007.

A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005. ISBN 0521837162.

F. Gustafsson and F. Gunnarsson. Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements. *Signal Processing Magazine, IEEE*, 22(4):41–53, July 2005. ISSN 1053-5888. doi: 10.1109/MSP.2005.1458284.

Z.J. Haas. A new routing protocol for the reconfigurable wireless networks. In *1997 IEEE 6th International Conference on Universal Personal Communications Record, 1997*, volume 2, 1997.

J. Harri, F. Filali, and C. Bonnet. Mobility models for vehicular ad hoc networks: a survey and taxonomy. *Communications Surveys & Tutorials, IEEE*, 11(4):19–41, 2009. ISSN 1553-877X.

IEEE. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999)*, pages C1–1184, 12 2007.

G. Kramer, I. Marić, and R.D. Yates. Cooperative communications. *Foundations and Trends® in Networking*, 1(3):271–425, 2006. ISSN 1554-057X.

P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S.S. Panwar. CoopMAC: A cooperative MAC for wireless LANs. *IEEE Journal on Selected Areas in Communications*, 25(2):340, 2007.

C. L. F. Mayorga, F. della Rosa, S. A. Wardana, G. Simone, M. C. N. Raynal, J. Figueiras, and S. Frattasi. Cooperative positioning techniques for mobile localization in 4g cellular networks. *Proceedings of the IEEE International Conference on Pervasive Services*, July 2007.

C. Mensing, S. Sand, M. Laaraiedh, B. Uguen, B. Denis, M. Garcia, J. Casajus, T. Pedersen, G. Steinboeck X. Yin, B.H. Fleury, S. Mayrargue, and D. Slock. Where d2.1 - performance assessment of hybrid data fusion and tracking algorithms. Technical report, ICT-217033 WHERE, 2008.

M. Musolesi and C. Mascolo. Mobility models for systems evaluation. a survey, 2009.

A. Nosratinia, T.E. Hunter, and A. Hedayat. Cooperative communication in wireless networks. *Communications Magazine, IEEE*, 42(10):74–80, 2004. ISSN 0163-6804.

C. Pedersen, T. Pedersen, C. Mensing, K. Papakonstantinou, T.-M. Oktem, D. Slock, T. Tenoux, M. Laaraiedh, B. Uguen, B. Denis, M. Garcia, M. A. Garcia, and J. Youssef. Where d2.3 - hybrid localization techniques. Technical report, ICT-217033 WHERE, 2010.

T.S. Rappaport. *Wireless communications*. Prentice Hall PTR, 2002. ISBN 0130422320.

A.H. Sayed, A. Tarighat, and N. Khajehnouri. Network-based wireless location: challenges faced in developing techniques for accurate wireless location information. *Signal Processing Magazine, IEEE*, 22(4):24–40, July 2005. ISSN 1053-5888. doi: 10.1109/MSP.2005.1458275.

Guolin Sun, Jie Chen, Wei Guo, and K.J.R. Liu. Signal processing techniques in network-aided positioning: a survey of state-of-the-art positioning designs. *Signal Processing Magazine, IEEE*, 22(4):12–23, July 2005. ISSN 1053-5888. doi: 10.1109/MSP.2005.1458273.

S. Valentin, H.S. Lichte, H. Karl, G. Vivier, S. Simoens, J. Vidal, and A. Agustin. Cooperative Wireless Networking Beyond Store-and-Forward. *Wireless Personal Communications*, 48(1):49–68, 2009. ISSN 0929-6212.

H. Wymeersch, J. Lien, and M. Z. Win. Cooperative localization in wireless networks. *Proceedings of the IEEE*, February 2009.

J. Yoon, M. Liu, and B. Noble. Random waypoint considered harmful. In *IEEE Societies INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications*, volume 2.

3

Distance-dependent Throughput Model for WiFi

This chapter describes the throughput model that is used to calculate the expected throughput for Wi-Fi networks throughout this thesis. Specifically, it is being used for a priori estimation of link qualities for optimal configuration choice in the chapters 5, 6, and 7. The model calculates the saturation throughput for a single Wi-Fi link between a source and a destination node, given the nodes' separation distance, transmit power, and modulation scheme. Further, the model is able to take into account a single interferer, if present. The throughput is calculated using a simple analytical model that takes into account the retransmission scheme of the IEEE 802.11 DCF in basic (acknowledged) mode, however assuming that no collisions occur.

3.1 Introduction

As outlined in chapter 1, the theme of this thesis is to investigate location based network optimizations, and under which circumstances they can help to improve throughput. Location information in itself does not directly describe throughput performance. In order to determine which configurations can enhance system performance, it is necessary to have a throughput model that can estimate the achievable performance of different configuration choices.

The network optimizations considered in chapters 5, 6, and 7, all require a model for judging the achievable performance of different configurations, thereby allowing to choose the most suited configuration. The requirements of each optimization are as follows:

- The location based relaying considered in Chapter 5, needs to determine the most suited relay node, depending on the locations of the destination node and the candidate relay nodes. The analysis in this chapter focuses on the impact of mobility, measurement collection delay, and inaccurate input information and does not consider cross-traffic or interfering nodes.
- The cross-layer optimization considered in Chapter 6 uses location information to enable simultaneous relay-to-destination transmissions, by interference aware selection of relay nodes and relay transmit power. The optimization algorithm needs to be able to make a priori estimation of the impact of interference and path loss for different choices of relay nodes and relay transmission power. This analysis also does not consider cross-traffic or interfering nodes, besides the interference caused by the simultaneous relay-to-destination transmission.
- The handover optimization considered in Chapter 7 uses location information and prediction of node movements to determine which networks to handover to. Here it is necessary to estimate in advance the achievable throughput from each network along the predicted movement trajectory, in order to facilitate planning of future handovers. The analysis considers only a single user node, and does not take into account cross-traffic or interfering nodes of any kind.

Since Wi-Fi networks are considered in the scenarios described in chapter 1, the model should take into account the functionality of Wi-Fi networks, namely the IEEE 802.11 protocol.

The IEEE 802.11 protocol specifies both PHY and MAC layer functionality, as described in chapter 2. There are different PHY layer standards available

(a,b,g,n), which have different properties with regard to for example the used frequency bands and bit rates. The proposed model can be used with any of these PHY standards, by using the constants defined in appendix A, that correspond to the considered technology, and by calculating the BER and FER according to the used modulation and coding scheme. For this thesis the constants corresponding to the 802.11a mode are used, since this is a widely used variant.

On the MAC layer, the model should take into account the impact of bit and frame errors due to low SNR/SINR, as well as the timing aspects of the IEEE 802.11 DCF. The DCF can use either basic acknowledged scheme for transmissions or the RTS/CTS scheme, which improves throughput in situations with much contention and a high risk of collisions. As cross-traffic and interfering nodes are not considered in the network optimization analyses, it is assumed that only the basic acknowledged DCF scheme is used. This scheme copes with transmission errors by retransmitting unacknowledged until the retransmission limit is reached. As poor link conditions can cause frame losses and thereby retransmissions, this feature should be modeled.

Based on the above descriptions, the following required properties have been identified:

- Degradation of throughput depending on distance, i.e., BER and Frame Error Rate (FER) impact of decreasing Signal to Noise Ratio (SNR).
- Throughput impact of a single interfering simultaneous transmission, i.e., BER and FER impact of decreasing Signal to Interference and Noise Ratio (SINR).
- BER model according to IEEE 802.11a PHY modulation scheme.
- Retransmission protocol for 802.11 MAC DCF basic acknowledged scheme.

The following features were identified as not needed:

- Not necessary to account for collisions and interference, since cross-traffic from nearby nodes is not considered.
- Not modeling RTS/CTS scheme in 802.11 DCF, since collisions are not considered.

3.2 Related Work

An accurate model of the IEEE 802.11 MAC DCF for basic (acknowledged) and RTS/CTS modes is presented in [Bianchi, 2000]. This widely applied Markov-chain based model however assumes that transmission errors are caused only by collisions caused by neighboring nodes. It is intended for cases where all nodes have similar good signal quality and are within the same collision domain and it is also based on the assumption that the network is saturated. Since the model does not model errors caused by low signal quality due to distance, the model is unsuited for use in the present work.

In several papers, such as [Malone et al., 2007] and [Cantieni et al., 2005] the authors have extended Bianchi's model to account for unsaturated network conditions, which is achieved by adding an extra idle state to the Markov chain. However, both mentioned papers have the assumption of an error-free channel, and they are therefore both unsuited for use in the present work.

There are several examples in the literature of work where the impact of channel induced errors on the throughput are taken into account. For example in [Dong and Varaiya, 2005] the authors have extended Bianchi's model to take into account frame losses due to transmission errors. This is done by introducing the probability of a backoff due to a transmission error, which is assumed to be independent of the collision error. The model takes a common BER as input to characterize channel quality. In [Ni et al., 2005] another extension of Bianchi's model is proposed, which allows to investigate the saturation throughput performance achieved at the MAC layer, in both congested and error-prone channels. The transmission error probability is however common for all nodes. This model also takes a common BER as input to characterize channel quality. Another approach also based on Bianchi's model, is presented in [Zheng et al., 2006], where the authors propose a model that takes into account both the incoming traffic loads and the effect of imperfect wireless channels, in which frame losses may occur due to bit transmission errors. However, control frames and data frame headers are assumed to be error free and the BER is common for all nodes. Finally in [Daneshgaran et al., 2008] the authors have extended Bianchi's model to take into account channel induced errors and the capture effect. Their results show that their model allows accurate derivation of the achievable throughput based on a multitude of system parameters for any of the different 802.11 technologies. The model takes a common SNR as input, which is used to calculate the BER.

Common for these examples of existing work is that they are based on Bianchi's model of the DCF. The original purpose of this model was to account

for the impact of contention and collision among stations in a WLAN and estimate the saturation throughput. The model required for the present work does however not need to account for contention and collision. Therefore a simpler model which only accounts for channel errors due to distance would be preferable. Even if future work should consider more detailed analyses in which cross-traffic is taken into account, it would be necessary for the considered model to account for different link qualities resulting from nodes' different locations. Specifically, different link qualities are typically compensated for by the 802.11 network adapter by adapting the used bit error rate, so that the best throughput is achieved.

Since a suited throughput model which satisfies the required features was not identified in existing work, the remaining parts of this chapter describes the proposed model.

3.3 Throughput Model

The proposed throughput model for Wi-Fi networks satisfies the features listed in section 3.1. The model corresponds to IEEE 802.11a as described in [IEEE, 2007], with the already discussed exception of no cross-traffic, and the exception that for simplicity, convolutional coding is not taken into account in the calculation of the FER. The latter is expected to result in conservative estimates of the FER.

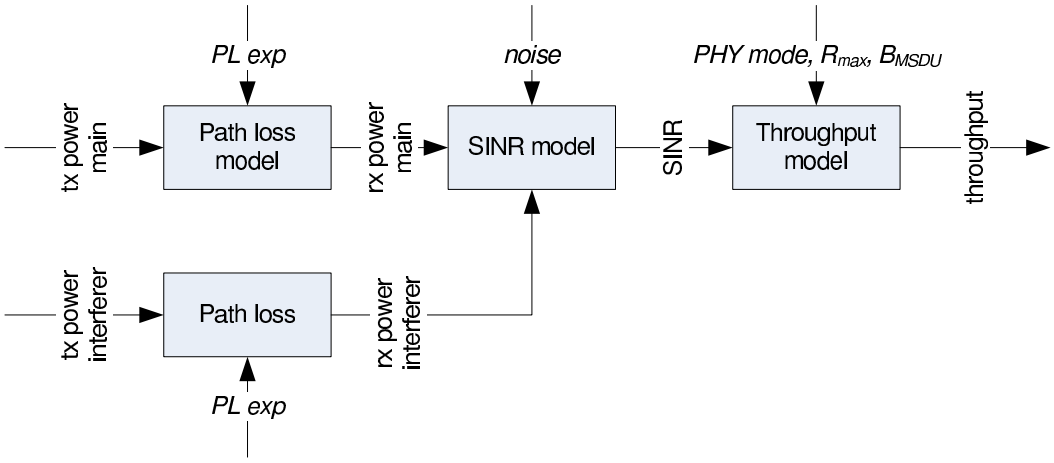


Figure 3.1: Structure of throughput model. Model parameters are italicized.

The throughput model is used to calculate the expected throughput for a transmission between network entities, given node positions and transmit

power levels. For calculating the throughput, the SNR (or SINR if interferers are present) and in turn average BER are first calculated. Based on the BER, the FER is calculated given different transmission parameters, and finally the throughput delivered by the MAC layer is calculated by taking into account the retransmission feature of the 802.11 MAC layer. In the following these calculations are given in details.

For the case with no interfering stations, the SNR is calculated as:

$$\gamma \ [dB] = 10 \ log_{10} \left(\frac{P_{\text{rx}}(d_{\text{rx-tx}}) \ [mW]}{N_0 \ [mW]} \right) \tag{3.1}$$

where $P_{\text{rx}}(d_{\text{rx-tx}})$ is the received power from the main transmitter at distance $d_{\text{rx-tx}}$ from the receiver and N_0 is the assumed noise floor. The received power is calculated using the path loss model in equation (2.1).

For the case of a single interfering station, the SINR is calculated as:

$$\gamma \ [dB] = 10 \ log_{10} \left(\frac{P_{\text{rx}}(d_{\text{rx-tx}}) \ [mW]}{P_{\text{rx}}(d_{\text{rx-interf}}) \ [mW] + N_0 \ [mW]} \right) \tag{3.2}$$

where $P_{\text{rx}}(d_{\text{rx-interf}})$ is the power received from the interfering transmitter at distance $d_{\text{rx-interf}}$ from the receiver.

The SNR/SINR is then converted into average BER for the used modulation scheme (BPSK, QPSK, 16-QAM or 64-QAM for IEEE 802.11a) and fading environment using for example theoretical expressions from reference [Proakis, 1995]. In this thesis, we have used the matlab function `berfading` from the *Communications Toolbox*, which calculates the average BER for a specific modulation scheme given a Rayleigh or Ricean fading channel.

In the IEEE 802.11 MAC acknowledged mode, there are three different outcomes of a frame transmission, assuming that the PHY header was successfully transmitted. A successful reception of a frame (s) is possible if no bit errors occur. If bit errors occur, one of the following two outcomes are possible: failed during DATA (fd) and failed during ACK (fa).

Assuming a constant and independent BER denoted P_{b}, the outcomes have the following probabilities:

$$P_{\text{s}} = (1 - P_{\text{b}})^{N_{\text{data}}+N_{\text{ACK}}} \tag{3.3}$$

$$P_{\text{fd}} = 1 - (1 - P_{\text{b}})^{N_{\text{data}}} \tag{3.4}$$

$$P_{\text{fa}} = (1 - P_{\text{b}})^{N_{\text{data}}} \cdot (1 - (1 - P_{\text{b}})^{N_{\text{ACK}}}) \tag{3.5}$$

where N_{data} and N_{ACK} are the number of bits transmitted in data and ACK frames, respectively, as defined in Appendix A on page 173. Here it is assumed

that the ACK frame has the same BER as the data frame. This may not be the case if the source and destination use different transmit power levels or if there are nearby interferers.

Given the constants in Appendix A on page 173, we calculate the average time of successful and failed transmissions as:

$$T_{\mathrm{s}}(r) = \overline{T_{\mathrm{BO}}}(r) + T_{\mathrm{data}} + T_{\mathrm{SIFS}} + T_{\mathrm{ACK}} + T_{\mathrm{DIFS}} \tag{3.6}$$

$$\begin{aligned}\overline{T_{\mathrm{f}}(r)} = &\left(\overline{T_{\mathrm{BO}}}(r) + T_{\mathrm{data}} + T_{\mathrm{DIFS}}\right) \cdot P_{\mathrm{fdn}} + \\ &+ \overline{T_{\mathrm{BO}}}(r) + (T_{\mathrm{data}} + T_{\mathrm{SIFS}} + T_{\mathrm{ACK}} + T_{\mathrm{DIFS}}) \cdot P_{\mathrm{fan}}\end{aligned} \tag{3.7}$$

where, $P_{\mathrm{fdn}} = \frac{P_{\mathrm{fd}}}{P_{\mathrm{fd}}+P_{\mathrm{fa}}}$ and $P_{\mathrm{fan}} = \frac{P_{\mathrm{fa}}}{P_{\mathrm{fd}}+P_{\mathrm{fa}}}$ are normalization factors, and $\overline{T_{\mathrm{BO}}}(r)$ is the average back-off time, which depends on the number of the current retry attempt r (where $0 \leq r \leq R$, and R is the maximum number of retry attempts). Hence, also $T_{\mathrm{s}}(r)$ and $T_{\mathrm{f}}(r)$ depend on r. According to [IEEE, 2007], the contention window (CW) is a uniform Random Variable (RV) between $CW_{\mathrm{min}} = 15$ and $CW_{\mathrm{max}} = 1023$. For each consecutive retry the CW is set according to:

$$CW(r) = min(1023, 2^{4+r} - 1). \tag{3.8}$$

We assume the average waiting time due to back-off is:

$$\overline{T_{\mathrm{BO}}}(r) = T_{\mathrm{slot}} \cdot \frac{CW(r)}{2} \tag{3.9}$$

where T_{slot} is the slot time used in IEEE 802.11.

As we have now determined the time and probability of a single successful or failed transmission, we now derive the expected throughput delivered by the MAC layer service, when taking MAC layer retransmissions into account. In IEEE 802.11 the default maximum number of retransmission attempts, here denoted R is 7. After R attempts the frame transmission fails and an error will be returned from the MAC layer without delivering the payload. In this work we only consider the MAC throughput, which may be different from the throughput achieved by overlying transport protocols and applications, due to for example time-out mechanisms as used in TCP to judge when a segment has been lost and needs to be retransmitted [Kesselman and Mansour, 2005]. Let n be the retry number, and T_{tx} the RV representing the time spent on a transmission attempt:

$$T_{\mathrm{tx}}(n) = \begin{cases} T_{\mathrm{s}}(n) & \text{for} \quad n = 0 \\ \sum_{r=0}^{n} T_{\mathrm{f}}(r) + T_{\mathrm{s}}(n+1) & \text{for} \quad 0 \leq n \leq R-1 \\ \sum_{r=0}^{n} T_{\mathrm{f}}(r) + T_{\mathrm{s}}(n+1) & \text{for} \quad n = R, \text{ success} \\ \sum_{r=0}^{n+1} T_{\mathrm{f}}(r) & \text{for} \quad n = R, \text{ failure} \end{cases} \tag{3.10}$$

with outcome probabilities:

$$P_{\text{tx}}(n) = \begin{cases} (1-P_{\text{s}})^n \cdot P_{\text{s}} & \text{for} \quad 0 \leq n \leq R-1 \\ (1-P_{\text{s}})^n \cdot P_{\text{s}} & \text{for} \quad n = R, \text{ success} \\ (1-P_{\text{s}})^{n+1} & \text{for} \quad n = R, \text{ failure} . \end{cases} \tag{3.11}$$

From this, we can compute the expected value as:

$$E[T_{\text{tx}}] = \sum_{n=0}^{R-1} (T_{\text{tx}}(n) P_{\text{tx}}(n)) + T_{\text{tx}}^{\text{s}}(R) P_{\text{tx}}^{\text{s}}(R) + T_{\text{tx}}^{\text{f}}(R) P_{\text{tx}}^{\text{f}}(R) \tag{3.12}$$

where $T_{\text{tx}}^{\text{s}}(R)$, $P_{\text{tx}}^{\text{s}}(R)$, $T_{\text{tx}}^{\text{f}}(R)$, and $P_{\text{tx}}^{\text{f}}(R)$ are the transmission time per attempt and frame delivery probability for the successful and failed cases in equations (3.10) and (3.11). The probability of a successful MAC frame delivery is:

$$P_{\text{suc}} = 1 - (1 - P_{\text{s}})^{R+1}. \tag{3.13}$$

Notice however that in practical systems ACK frames may be transmitted with a different bit rate than the data frame. In this case, it would be necessary to rewrite equations (3.3), (3.4), and (3.5) to allow different bit error probabilities for the data and ACK frames.

As throughput is given by $\frac{\text{Delivered data}}{\text{Transmission time}}$, the throughput can be calculated as:

$$S = \frac{P_{\text{suc}} \cdot B_{\text{MSDU}}}{E[T_{\text{tx}}]} \tag{3.14}$$

where B_{MSDU} is the MAC payload size is given in octets.

3.4 Results and Discussion

In this section a few example results of the described throughput model are shown. Also, for verification of the throughput model, a custom matlab simulation of the IEEE 802.11 back-off mechanism has been created. This is basically a simulation of the model, where the transmission and reception of the individual data and ACK frames over a link are simulated. The outcome of each frame transmission is stochastically determined based on the data frame size and the BER, and the BER level depends on the SNR level for the particular transmission. These simulation results are shown as circles in the following plots.

The plot in Figure 3.2 shows the resulting FER for different BER levels. The simulation results are based on 1 000 000 simulation runs. Some of the circles are not aligned perfectly with the analytic result in the lower part of the plot where the FER is low. For these low FER levels a higher number of simulation runs are needed for perfectly aligned results.

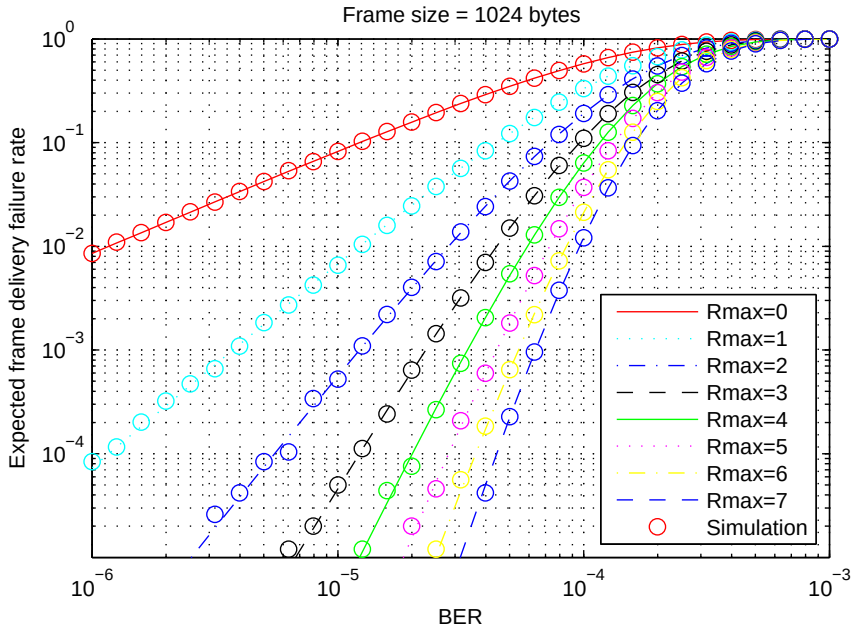


Figure 3.2: Expected FER $(1 - P_{\text{suc}})$ for 6 Mbit/s and $B_{\text{MSDU}} = 1024$ bytes.

The following results in Figure 3.3 shows the corresponding results of the expected time of each transmission attempt, at different BER levels.

In Figure 3.4 the previous two results are combined to give the throughput. The plot includes different curves for each value of R_{max}, however in practice there is not a very big difference between $R_{\text{max}} = 0$ and $R_{\text{max}} = 7$, which could suggest that a practical system need not to retransmit. However, in the cases where R_{max} is low, an overlying transport protocol or application is responsible for handling any frame losses that may occur, which is typically less efficient than MAC retransmissions.

Finally, in Figure 3.5 the distance throughput relationship is exemplified for some of the different modulation schemes available in IEEE 802.11a. The plot is generated using the throughput model for the scenario used in chapter 7, where there are no interferers present. The model results are clearly consistent with the simulation results. An interesting future work item would be to

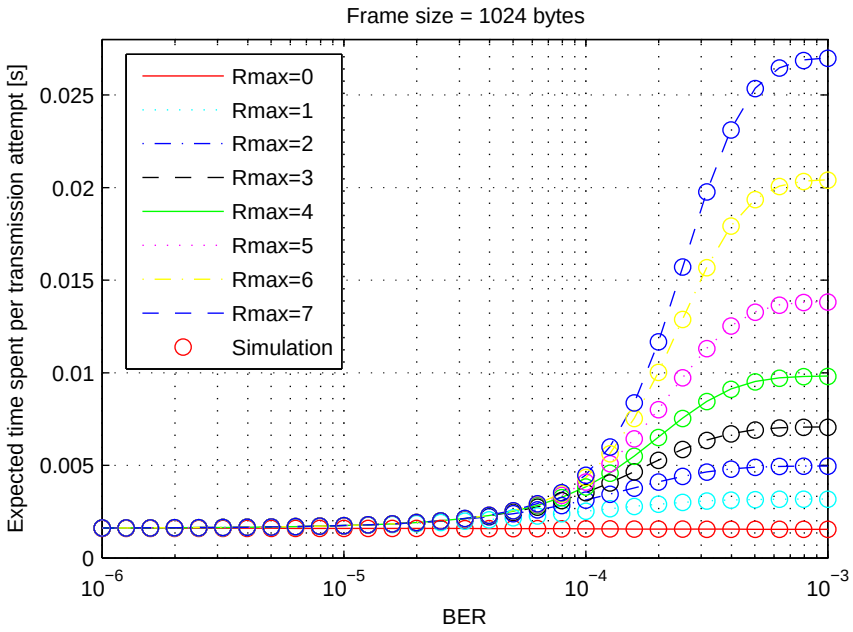


Figure 3.3: Expected transmission time ($E[T_{\text{tx}}]$) for 6 Mbit/s and $B_{\text{MSDU}} = 1024$ bytes.

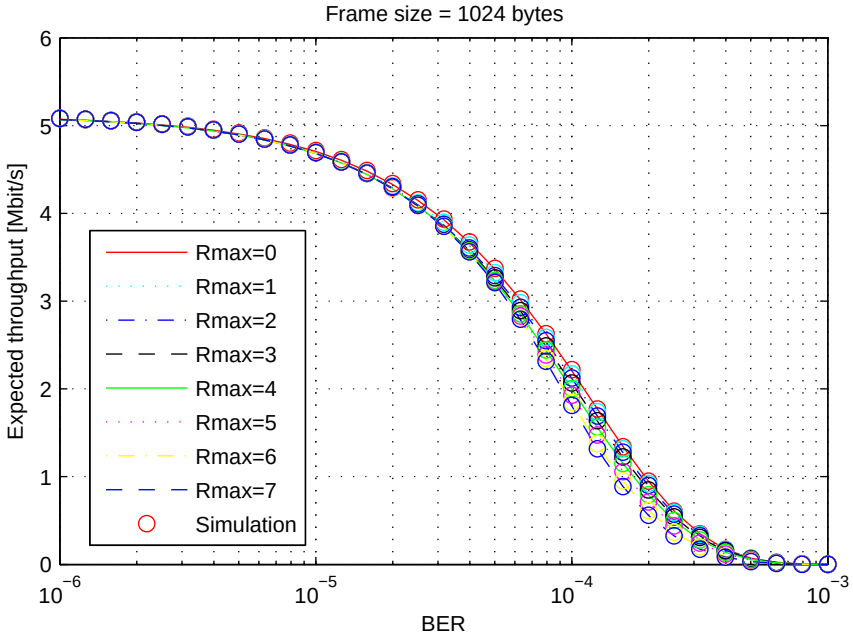


Figure 3.4: Expected average throughput for 6 Mbit/s and $B_{\text{MSDU}} = 1024$ bytes.

validate the model assumptions using measurements or commonly used simulation tools such as ns-2 or OMNET++. Further, this would make it easier to compare the model to other models in the literature.

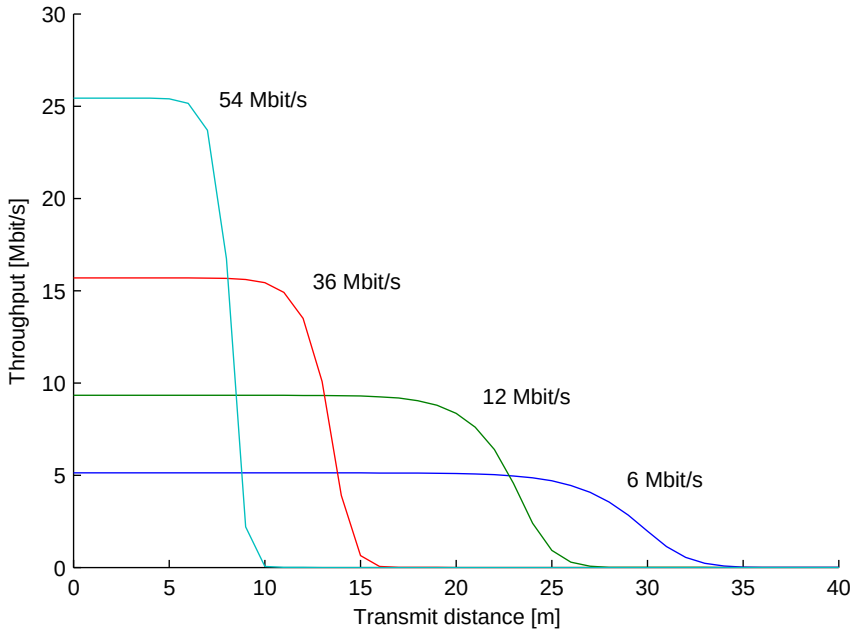


Figure 3.5: Expected average throughput for 6, 12, 36, and 54 Mbit/s transmission modes, generated using the throughput model. Here it is assumed that $R_{\max} = 7$ and $B_{\text{MSDU}} = 1024$ bytes and that the scenario parameters are similar to the scenario considered in chapter 7, which is specified by the parameters in Table 7.1 on page 149.

3.5 Conclusion

In this chapter a simple throughput model of IEEE 802.11a based Wi-Fi networks has been presented. The model focuses on the impact of distance on the achieved throughput, and is intended to be used for guiding the choice of relaying and handover parameters for the location based network optimizations considered in chapters 5, 6, and 7.

The model calculates the saturation throughput for a single Wi-Fi link between a source and a destination node. The model inputs are the nodes' separation distance, transmit power, and modulation scheme. The model is able to take into account a single interferer, if present. First, the BER is calculated according to the 802.11a modulation schemes, however for simplicity, the

convolutional coding and Viterbi decoding of 802.11a has not been taken into account, which means that the calculated throughput is lower than what it would be if these error correction properties were taken into account. In [Qiao et al., 2002] an upper bound is used to estimate the packet error probability of 802.11a with Viterbi decoding, however only for the AWGN channel. An estimate for the Ricean channel could be found using simulations, however this was not considered in this work.

Hereafter the Frame Error Rate is calculated under the assumption of independent bit errors. The saturation throughput is then estimated using a simple analytical model that takes into account the retransmission scheme of the IEEE 802.11 DCF in basic (acknowledged) mode. The model is intended for a priori estimation of the link quality, in order to choose the best configuration in the network optimization algorithms considered in chapters 5, 6, and 7. As the analyses of these algorithms do not consider cross-traffic, the model does not need to model the impact of collisions on the 802.11 DCF. Therefore the channel is always assumed to be idle when a transmission is started. The model is much simpler than existing models based in Bianchi's model in [Bianchi, 2000], such as [Daneshgaran et al., 2008, Zheng et al., 2006, Ni et al., 2005, Dong and Varaiya, 2005], as these models account for the impact of both channel errors and cross-traffic. Finally, simulations of the 802.11 DCF retransmission scheme have been used to verify the proposed model. An interesting future work item would be to validate the model results using measurements or commonly used simulation tools such as ns-2 or OMNET++.

References

G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *Selected Areas in Communications, IEEE Journal on*, 18(3):535–547, 2000.

G.R. Cantieni, Q. Ni, C. Barakat, and T. Turletti. Performance analysis under finite load and improvements for multirate 802.11. *Computer Communications*, 28(10):1095–1109, 2005. ISSN 0140-3664.

F. Daneshgaran, M. Laddomada, F. Mesiti, and M. Mondin. Unsaturated throughput analysis of IEEE 802.11 in presence of non ideal transmission channel and capture effects. *Wireless Communications, IEEE Transactions on*, 7(4):1276–1286, 2008. ISSN 1536-1276.

X.J. Dong and P. Varaiya. Saturation throughput analysis of ieee 802.11 wireless lans for a lossy channel. *Communications Letters, IEEE*, 9(2):100 – 102, 2005. ISSN 1089-7798. doi: 10.1109/LCOMM.2005.02011.

IEEE. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999)*, pages C1–1184, 12 2007.

A. Kesselman and Y. Mansour. Optimizing TCP retransmission timeout. In *ICN'05: Proc. of The 4th International Conference on Networking*, pages 133–140. Citeseer, 2005.

D. Malone, K. Duffy, and D. Leith. Modeling the 802.11 distributed coordination function in nonsaturated heterogeneous conditions. *Networking, IEEE/ACM Transactions on*, 15(1):159–172, 2007. ISSN 1063-6692.

Q. Ni, T. Li, T. Turletti, and Y. Xiao. Saturation throughput analysis of error-prone 802.11 wireless networks. *Wireless Communications and Mobile Computing*, 5(8):945–956, 2005. ISSN 1530-8677.

J.G. Proakis. Digital Communications Third Edition. *McGrawHill Inc, New York, USA*, 1995.

D. Qiao, S. Choi, and K.G. Shin. Goodput analysis and link adaptation for IEEE 802.11 a wireless LANs. *IEEE transactions on Mobile Computing*, pages 278–292, 2002. ISSN 1536-1233.

Yu Zheng, Kejie Lu, Dapeng Wu, and Yuguang Fang. Performance analysis of ieee 802.11 dcf in imperfect channels. *Vehicular Technology, IEEE Transactions on*, 55(5):1648 –1656, 2006. ISSN 0018-9545. doi: 10.1109/TVT.2006.878606.

4

Realistic Communication Constraints for Localization

The purpose of the contribution presented in this chapter is to investigate the achievable localization accuracy of a localization system for wireless mobile networks, as well as to investigate the impact of realistic measurement obtainment and collection. A specific use case system is considered that uses UWB for ranging, which is modeled using empirical measurements. These ranging measurements are collected through a Wi-Fi infrastructure and processed centrally. The analysis takes into account delays and frame losses in the measurement collection process, which allows to study which impact this collection has on the localization accuracy. The contribution is a part of a joint work with Christian Mensing from DLR in Germany. This chapter presents selected parts of this work. See the original paper in [Mensing and Nielsen, 2010] for further details on the localization aspects.

4.1 Introduction

When evaluating the performance of localization algorithms, it is common practice to assume that measurements are sampled periodically and that they are immediately available for processing. For localization of non-mobile users, this assumption is fine. However, for tracking of mobile users, the delay from the time it takes to obtain and collect the necessary measurements could in principle have an influence on the achieved accuracy. In cooperative localization approaches, where additional measurements need to be collected from P2P links, the influence of the collection delay could be even more pronounced. The aim of this chapter is therefore to study the impact of realistic communication constraints on the localization accuracy, by taking into account the measurement obtainment and collection delays. Also, from a networking point of view, it is interesting to study what the collection delays are and how much overhead is generated in the Wi-Fi network. This is especially interesting for Cooperative Positioning (CP), where also P2P measurements are taken into account, in addition to the measurements from links between anchor nodes and mobile nodes that are considered in conventional positioning.

There are basically two ways to realize CP (as described in section 2.3):

1. Using a centralized approach as in [Mayorga et al., 2007, Frattasi, 2007], where all measurements are collected in one central localization server, which is responsible for jointly processing the measurements to compute the needed location estimates. These can then be exploited in the network (e.g., for optimization of communication functions) or be sent back to the mobile devices. This approach is well suited for infrastructure based systems, where the central localization server can always be reached.

2. The other approach is to use distributed algorithms based on for example Bayesian inference [Wymeersch et al., 2009], where the individual mobile devices are responsible for computing their own location, based on information they share only with their surrounding neighbors via P2P links, as in [Wymeersch et al., 2009, Chan and So, 2009]. Since this type of algorithms does not rely on a central entity for processing, such algorithms are well suited for ad hoc networks, such as for example wireless sensor networks.

Since the scenarios considered in this work focus on infrastructure connectivity, the centralized localization approach with a central localization server is considered in this chapter.

Specifically, in this chapter, we investigate the performance of conventional and cooperative centralized positioning schemes under realistic communication constraints and measurement models from both the positioning and the networking perspective. The centralized infrastructure is based on the large scale scenario presented in Figure 1.1, however with only the Wi-Fi part being used.

This study is based on the assumption that ultra-wideband (UWB) technology is used for ranging. UWB is particularly suited for ranging [Gezici et al., 2005], since the wide frequency range makes it more possibly that at least some radio waves will go through or around obstacles. Further, the wide frequency range allows for a high time resolution, which is useful for time-based ranging as TOA or TDOA.

The ranging is realized by UWB TOA measurements, and it is assumed that all Anchor Nodes (ANs) and Mobile Stations (MSs) are equipped with UWB radios dedicated for ranging. The Wi-Fi infrastructure is used for collecting the measurements between the ANs and the MSs as well as the P2P measurements between MSs. Mobility of the users is exploited by application of a tracking algorithm based on Extended Kalman Filters (EKFs). Hence, simulation results will provide a realistic assessment of centralized CP in a high-mobility environment.

4.2 System Model

There are two kinds of entities in the considered system. The ANs that are UWB anchor nodes are located at fixed, perfectly known locations. The second type of entity is the MS, which can move around in two dimensions. Figure 4.1 shows an example of a system with three ANs and two MSs. In this work a Wi-Fi based network is used for collecting ranging measurements. For simplicity, it is assumed that all anchor nodes are also acting as Wi-Fi access points. Notice that in this chapter the ANs are represented by the symbol of a base stations, while they are in fact acting as UWB anchor nodes as well as Wi-Fi access points.

The arrows between the entities show the different ranges (r) or distances that are measurable in the system. The range between MSs, in the example denoted as $r_{\mathrm{MS1,MS2}}$, is only used by the cooperate positioning algorithm, not for the conventional positioning algorithm.

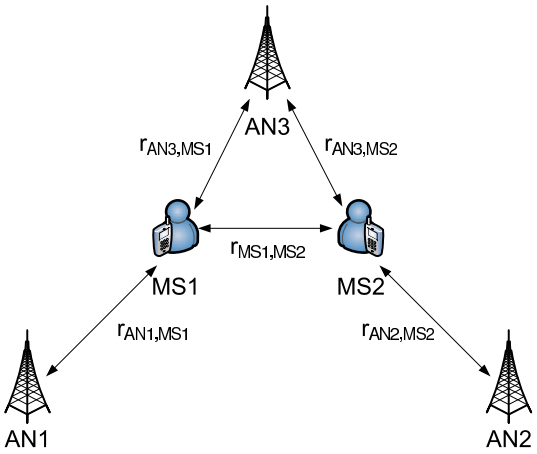


Figure 4.1: Example system with three ANs and two MSs.

For each range between a pair of entities, the following measurement error model is used to generate measurements:

$$\hat{r} = r + \mathbf{b} + \mathbf{n} \tag{4.1}$$

where $\hat{r}$ is the estimated distance, r is the actual distance, and the terms $\mathbf{b}$ and $\mathbf{n}$ are stochastic variables accounting for range estimation bias and noise, respectively.

For generating the ranging errors, we exploit the UWB device measurements performed within the WHERE project and the derived models from that. We make use of a preliminary version of the models presented in [Pezzin and Vazquez, 2010]. In summary, the devices use a 2 GHz bandwidth centered around 4 GHz, and a peak transmit power of 2 dBm. The full set of parameters for the UWB devices are available in the mentioned paper. The applied version of the model includes bias and residual noise conditioned on distance, orientation, and LOS/NLOS status of the connection. The average standard deviation of noise and the average bias are depicted in Figure 4.2 over the distance for LOS and NLOS conditions. We assume that the MS-MS connections are always LOS, whereas the MS-AN connections are NLOS in 50% of the cases.

4.3 Measurement Collection

For both the conventional and the cooperative approaches for localization that are considered in this work we have defined protocols that are responsible

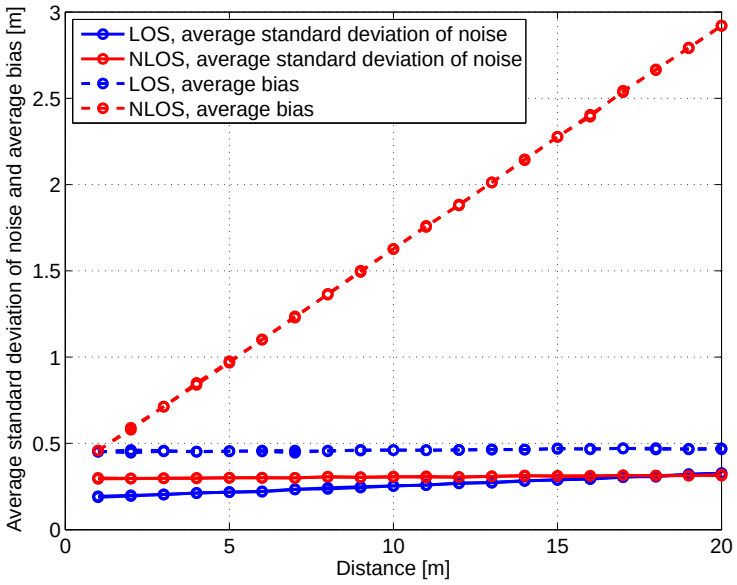


Figure 4.2: Average ranging error model parameters vs. distance

for collection of measurements and provision of a location estimate. In the following we describe these protocols. It is assumed that the location estimate is needed by an application at the MS, which polls the location every μ_{loc} seconds.

4.3.1 Device-based Conventional Localization Case

In this case, the localization algorithm uses only measurements from the MS-AN links as sketched in Figure 4.3 which shows an example scenario with 4 ANs. As the MS holds all measurements necessary to compute the location estimate, we assume the localization/tracking algorithm is run in the MS.

It is assumed that ranging measurements are obtained towards each of the four ANs periodically, each with interval $\mu_{\text{MS-AN}}$. Since the ranging measurements are obtained directly in the MS, the only additional factor that contributes to the localization delay is the application location request interval μ_{loc}.

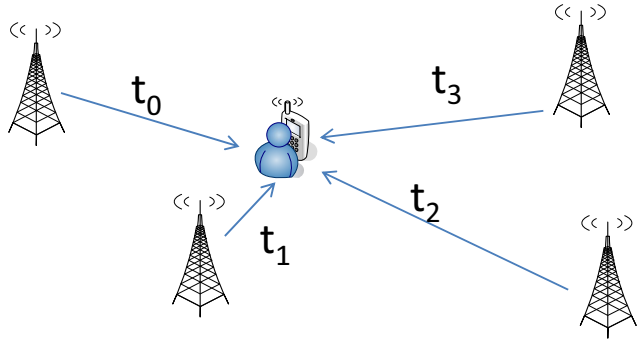


Figure 4.3: Message flow in device-based conventional positioning.

4.3.2 Centralized Cooperative Localization Case

In addition to MS-AN ranging measurements, the cooperative localization algorithms uses MS-MS ranging measurements and centralized computation of location estimates. In order to realize the collection of both types of measurements, as well as send back the location estimate to the MS, the message flow sketched in Figure 4.4 is used.

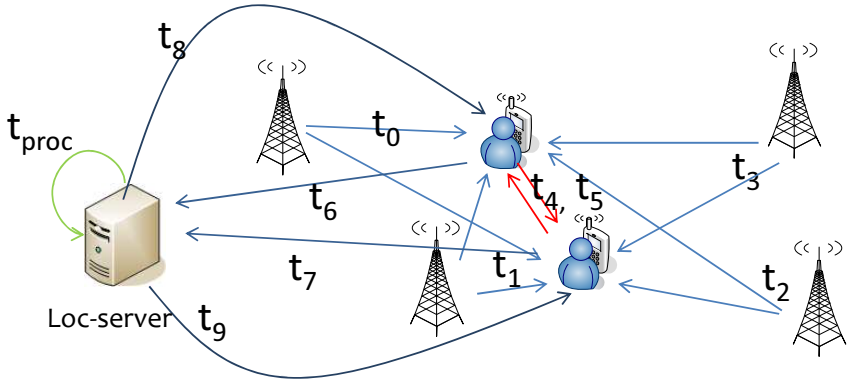


Figure 4.4: Message flow in centralized cooperative positioning. Ranging measurements are obtained in the MS by performing ranging towards ANs and other MSs. The obtained measurements are then collected in the localization server for processing. Finally, the location estimate is sent back to the MS.

In order to show the message flow more clearly, we consider the subflows individually in the following.

Like the conventional algorithms, the cooperative algorithms rely on periodically obtained ranging measurements (every $\mu_{\text{MS-AN}}$ seconds) for MS-AN measurements. Figure 4.5 shows how ranging measurements are first obtained between the MS and AN. Hereafter a measurement packet, which contains the ranging measurement, is sent to the nearest AN and thereafter to the localization server, which is assumed to be connected to the AN by a wired infrastructure. Here it would be possible to bundle measurements from the ANs and send them in one packet to the localization server. However, here we have decided to use the flow that is most similar to the conventional case.

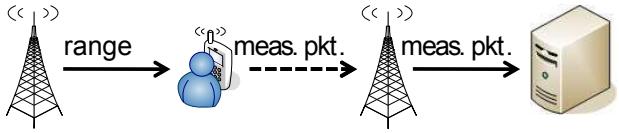


Figure 4.5: Message flow for MS-AN ranging measurements in cooperative localization.

In addition to MS-AN measurements, the cooperative algorithms rely on MS-MS measurements, also called P2P measurements. The flow of messages is shown in Figure 4.6. Whenever an MS senses another MS within D_{coop} meters, a P2P ranging measurement is made and sent to the localization server through the nearest AN. However, to reduce the amount of measurement packets being sent, P2P measurements are buffered and sent in a bundle every μ_{coop} seconds. As with the MS-AN measurements, the AN is assumed to be connected to the localization server by a wired infrastructure.

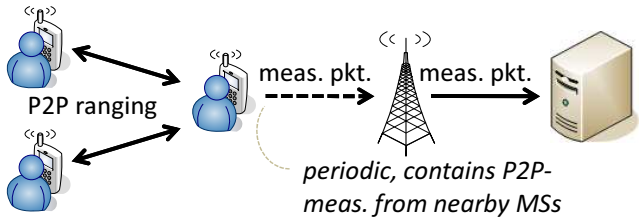


Figure 4.6: Message flow for P2P measurements in cooperative localization.

Having both MS-AN and MS-MS measurements at the localization server, we now need to provide the calculated position estimate to the MS. This is done by periodically unicasting a message with the current location estimate of an MS to that MS, as sketched in Figure 4.7.

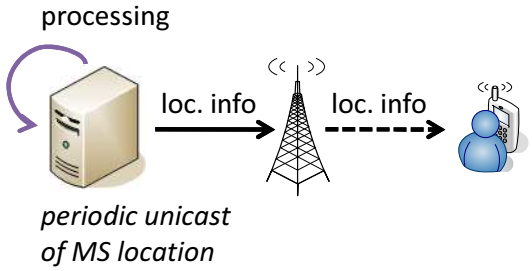


Figure 4.7: Message flow for location info message in cooperative localization.

A collected view of the flows in an example scenario is given in Figure 4.8 in the form of a sequence chart.

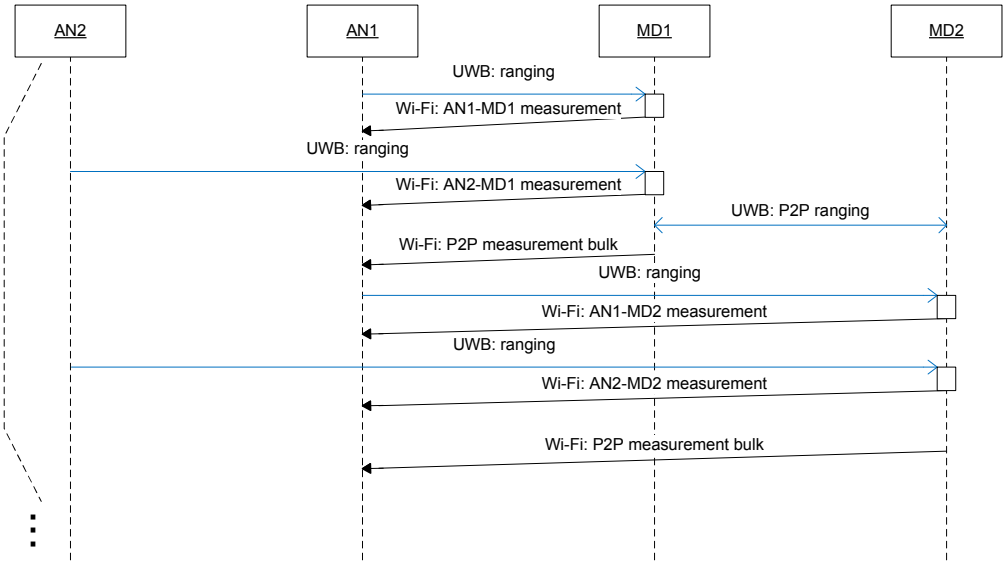


Figure 4.8: Sequence chart for an example scenario with 2 ANs and 2 MSs. In this example AN1 is closer to both MSs than AN2, and the MSs therefore send their obtained measurements through the AP in AN1. The UWB ranging measurement procedure is depicted with blue arrows.

Having described how measurements are collected at the server, the following describes how the localization algorithms use the measurements.

4.4 Localization Algorithms

For localization and tracking, two different algorithms have been used in this work. As these algorithms are not a part of the contributions in this work, please see the original paper [Mensing and Nielsen, 2010] for a detailed presentation. In the following, the key features are outlined.

The considered algorithms are using the ranging measurements that have been collected using the collection schemes presented in the previous section. The localization algorithms are executed periodically, every 1 second, and use the measurements that have been collected since the last execution of the algorithms.

Static Solution

The static solution is calculated independently for each time sample as a weighted non-linear Least Squares (LS) of the available measurements. Generally, there exists no closed form solution to this non-linear $2 \cdot N_{\mathrm{MS}}$-dimensional optimization problem, where N_{MS} is the number of MSs, which means that an iterative solution is necessary. The problem is solved using the Gauss-Newton (GN) algorithm, which however needs a good initial location estimate for fast convergence and good estimates. The initial solution used by the GN algorithm in this work is the geometric mean value of the involved ANs.

Notice that this algorithm uses only the available measurements for location estimation, it does not take advantage of previous location estimates.

Extended Kalman Filter Solution

The Extended Kalman Filter (EKF) solution is different from the static solution in the sense that it calculates updated position estimates in every time-step by considering the history of movement and new measurements according to the Bayesian philosophy. Also the EKF is able to take advantage of the fact that users move along certain trajectories, where each new position is strongly correlated to the previous.

The state space used for the EKF is $4 \cdot N_{\mathrm{MS}}$-dimensional, since it consists of two-dimensional positions and two-dimensional velocities for each considered MS. Further, the EKF takes as input the vector $\hat{r}[k]$, which contains all the available ranging measurements for the kth time step. Notice that this vector changes between time steps, since it depends on the available measurements.

The evolution model used for the EKF includes a priori information about the MSs movements in order to cope with missing measurements. Even the situation that no AN or other MS is visible for a certain time can be handled by this approach. In that situation, the movement model compensates the missing measurements.

Calculation of the different MSs location estimates are decoupled in the EKF, i.e., it is assumed that they move independently of each other.

The EKF is initialized (at $k = 0$) by the result from of static solution at that time. The used state vectors contain both positions and velocities of the MSs.

4.5 Evaluation Methodology

The considered localization algorithms have been evaluated with realistic communication constraints in a 4-step process as sketched in Figure 4.9.

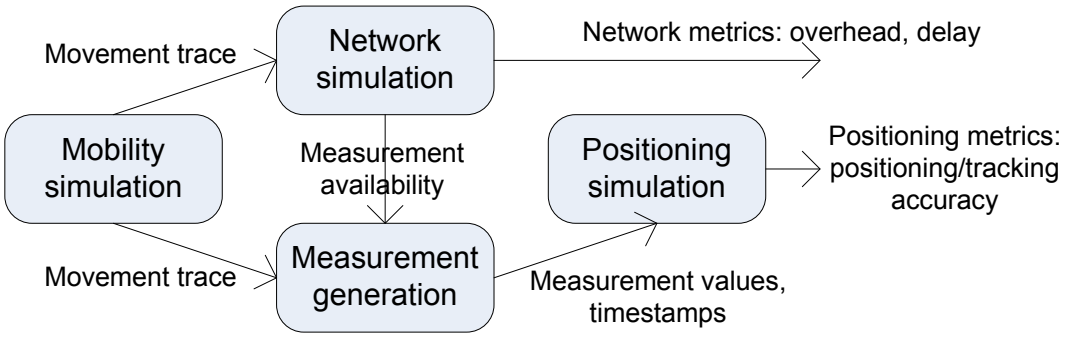


Figure 4.9: Simulation overview.

The mobility simulation is run initially, which results in a trace file that describes the AN positions and MS movements according to the random waypoint group mobility model described in section 4.5.1. This mobility trace is then used as a basis for simulating the Wi-Fi network that is used for collection of ranging measurements (both MS-AN and MS-MS) and provisioning of location estimates from the localization server to the MSs. It is assumed that the network is based on IEEE 802.11a.

The network simulation is realized using ns-2[1] based on the mobility trace and the scenario specific parameters listed in Table 4.2. We use the *802.11ext* module in ns-2. Custom agents are installed on the simulation nodes so that the ranging measurement collection flow is simulated as specified in section 4.3. The purpose of the simulation is to get realistic measurements of the time it takes to collect ranging measurements and provide location estimates, as well as the MSs ability to reach the localization server through the Wi-Fi infrastructure. Therefore we use the ns-2 simulation, which accurately simulates the MAC behavior of IEEE 802.11, which means that delays and frame losses caused by factors such as collisions, retransmissions, backoff, and congestion are taken into account. The ns-2 simulation was parameterized according to Table 4.2. Further, this ns-2 version includes a Nakagami fading model, which can approximate a Ricean fading environment, by using the model parameters $\Gamma = n$ and $m = \frac{(K+1)^2}{2K+1}$, where K is the Ricean K-factor.

Table 4.1 shows the sizes of the used messages. We have made the following assumptions regarding the used messages. The MS-AN ranging measurement is a 802.11 data frame with a payload consisting of the MAC id (6 bytes) of the AP and the estimated range (2 bytes). The P2P measurement bulk message size depends on the number of MS neighbors in range. It is also a data frame where the payload is a 6 bytes MAC id and 2 bytes ranging value for each neighbor MS. Finally, the location information message is a data frame with the node coordinates (x,y) encoded with 8 bytes each. The duration of the different frames can be calculated using Appendix A.

Message type	**MSDU size (octets)**
MS-AN measurement	8
P2P measurement bulk	$(6+2) \cdot N_{\text{MS in range}}$
Location information	16

Table 4.1: Message types

The output of the network simulation is first the network-related performance metrics, and secondly this block also delivers a trace file specifying time stamps for when measurements are obtained and have been collected in the central localization server, according to the collection protocol.

The measurement generation block uses the network simulation trace file in combination with the mobility trace, to generate the actual measurement

[1]The ns-2 simulation is based on [Chen et al., 2007], which has been updated with the author's patch from June 5th, 2009.

values for the MS-AN and MS-MS links. The measurements are generated using the models described in section 4.2. Notice that the measurement values are generated so that they correspond to the node positions at the time the ranging measurements were performed.

The positioning simulation is run as the last step and positioning metrics are computed for the considered conventional and cooperative localization algorithms.

The considered performance metric for the localization and tracking is the cdf of the absolute two-dimensional position error $\varepsilon_{\text{error}}$, i.e.,

$$\text{CDF}\left(\varepsilon_{\text{error}}\right) = \text{Prob}\left(\left\|\hat{\boldsymbol{x}} - \boldsymbol{x}\right\|_2 \leq \varepsilon_{\text{error}}\right), \tag{4.2}$$

where it was averaged over all MSs in the scenario and several noise realizations, i.e., multiple independent realizations of the measurement generation and positioning simulation blocks in Figure 4.9. We further assume that the MS-MS connections are always LOS, whereas the MS-AN connections are NLOS in 50 % of the cases.

4.5.1 Group Mobility Model

A variation of the random waypoint that mimics group mobility is used in this work. In each group of nodes, one of the nodes acts as the reference node. For this node a waypoint and speed is chosen as usual for the random waypoint model (see [Raspopoulos et al., 2010]). For the remaining nodes in the group the same speed is used and a their waypoints are chosen, so that they are randomly placed within D_{spread} of the reference node's waypoint. An example of the resulting mobility tracks is shown in 4.10 in a $100x100m^2$ scenario. In this example there are 6 groups with 4 nodes in each group, shown with a unique color for each group.

4.6 Results and Discussion

The simulation results are presented in this section in two parts: first the communication related results and secondly the localization/tracking results. The positioning algorithms are evaluated both with perfect communication, i.e., error-free and instantaneous exchange of all required information, and secondly with realistic communication constraint. The baseline simulation parameters are shown in Table 4.2.

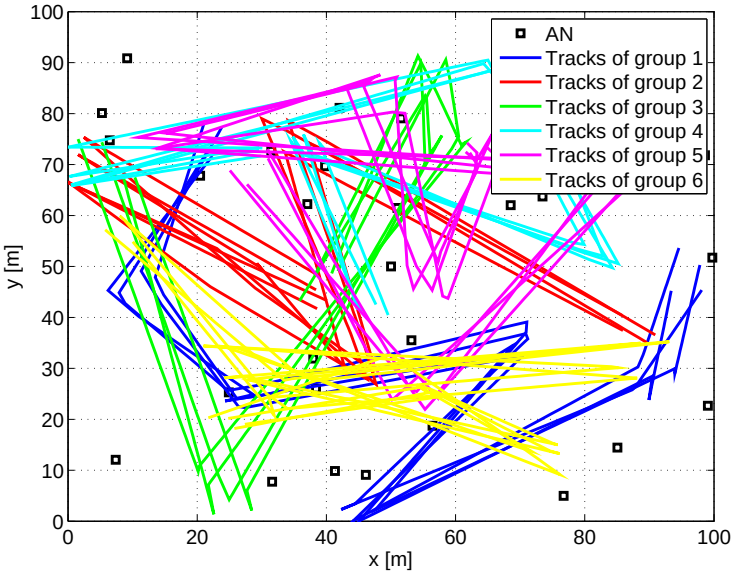


Figure 4.10: Group mobility simulation example.

4.6.1 Communication Part

Figure 4.11 shows the average localization delays for the conventional measurement collection and for the two types of measurements in the cooperative measurement collection. The localization delay is the time it takes from a ranging measurement is produced in an MS, until the polling application on the MS has an updated location estimate. The delay for the conventional collection protocol does not change, since its delay only depends on the MS-AN ranging interval $\mu_{\text{MS-AN}}$ and the polling interval of the application μ_{loc}. On the other hand, the delay of the cooperative collection protocol seems to increase slightly with the increase of the number of ANs.

With a mobile speed of 2 m/s and delays that are in the range between 0.5 to 1.4 seconds on average, the deviation caused by mobility and delay would be around 1 - 3 meters. However, since the ranging measurements are inaccurate due to noise and bias, and since the localization algorithms are using many different measurements for location estimation, the actual localization error is more complex to estimate.

Having presented the behavior of the network when considering realistic communications constraints, we now focus our attention on the positioning algorithms and how the realistic communications constraints affect them.

Simulation parameters	**Value**
Time	100 s
Size	$100x100m^2$
Number of ANs (N_{AN})	30
Mobility model parameters	**Value**
Number of MS groups (N_{groups})	6
Number of MSs per group ($N_{\text{MS/group}}$)	4
Max spread relative to ref. MS in group (D_{spread})	20 m
Movement speed ($\|\mathbf{v}\|$)	2 m/s
Protocol parameters	**Value**
AN ranging interval ($\mu_{\text{MS-AN}}$)	1 s
P2P ranging interval (μ_{coop})	1 s
P2P ranging distance (D_{coop})	20 m
Location information update interval ($\mu_{\text{loc-info}}$)	1 s
MS application request interval (μ_{loc})	1 s
Localization processing time (μ_{proc})	0.1 s
ns-2 PHY parameters	**Value**
Path loss exponent (n)	2.9
Rician K-factor (K)	6
Transmit power (P_{tx})	5 mW
802.11a PHY mode	6 Mbit/s, BPSK
Bandwidth	20 MHz
Frequency	5.18 GHz
Carrier Sense Threshold	−92 dBm
Noise floor	−106 dBm

Table 4.2: Simulation parameters

4.6.2 Positioning Part

Figure 4.12 shows the CDF for conventional (non-cooperative) and cooperative positioning for both static solution and tracking with EKF. We observe that for the static solution more than 10 % of the MSs cannot be localized (e.g., due to limited access to ANs or bad geometric conditions). This can be reduced by application of the EKF resulting in an error being smaller than 10 m in 90 % of the cases. If we allow cooperation between the MSs this can further be improved to around 3 m.

In addition to these results, Figure 4.13 shows also the results with realistic communications constraints. Here, we observe that the accuracy is decreased

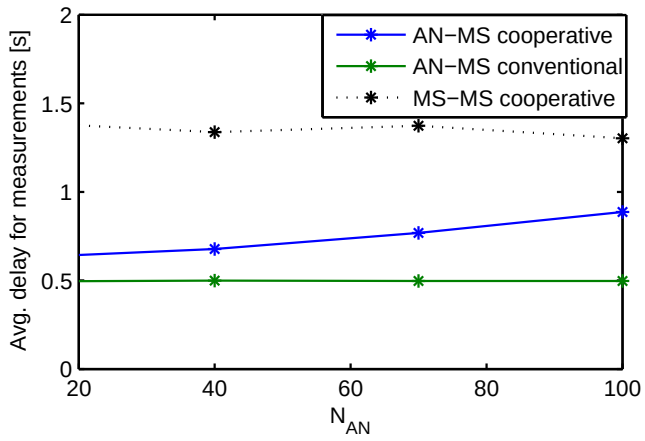


Figure 4.11: Average localization delay for varying number of ANs.

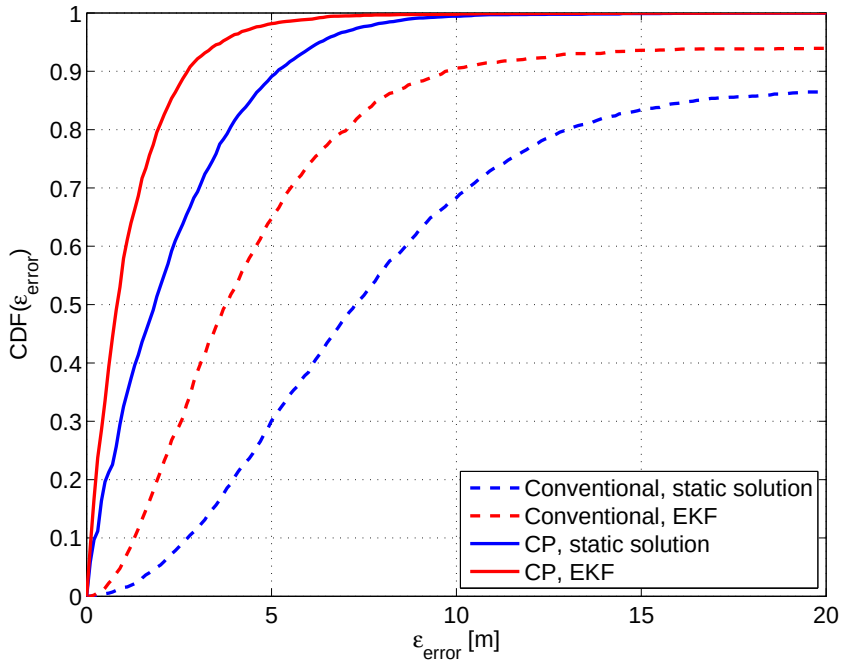


Figure 4.12: Conventional vs. cooperative positioning using static solution and EKF.

by 1 m in the conventional schemes, whereas it is reduced by around 2 m and 3 m for CP using static solution and EKF, respectively. As expected, the gap between realistic and perfect communications is higher for the CP scheme

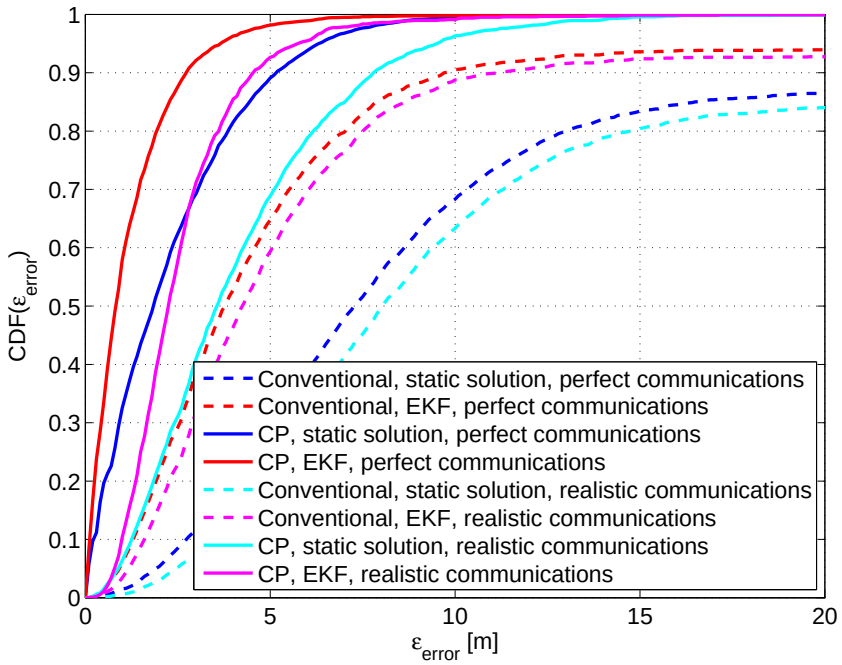


Figure 4.13: Conventional vs. cooperative positioning using static solution and EKF with realistic communications constraints.

compared to the conventional approach. Nevertheless, assuming CP and an EKF the 90 %-error is still below 5 m.

To evaluate the dependency on the MS-MS connectivity, in Figure 4.14 the number of MSs per group is varied. Note that an increased number of MSs per group automatically results in an increased overall number of MSs N_{MS} since the number of groups is kept constant. We observe that with only one MS per group no noteworthy gains can be achieved by CP compared to the conventional approach. Reason for that is that the connectivity between the groups is only limited. If we increase the number of MSs per group, e.g., to 10, cooperation can be exploited within each group and we achieve an 90 %-error of around 4 m in this scenario. If we increase it further to 20, it can be seen that the performance drops down rapidly, and in average around 12 % of the MSs cannot be localized. This could be explained by an increased communications overhead for performing CP with the $N_{\mathrm{MS}} = 120$ MSs and the resulting latency due to congestion.

Figure 4.15 depicts the dependency on the MS-AN connectivity. For a low number of ANs in the scenario (e.g., 10), several MSs cannot determine their position. In that situation also the cooperation gain is restricted since overall

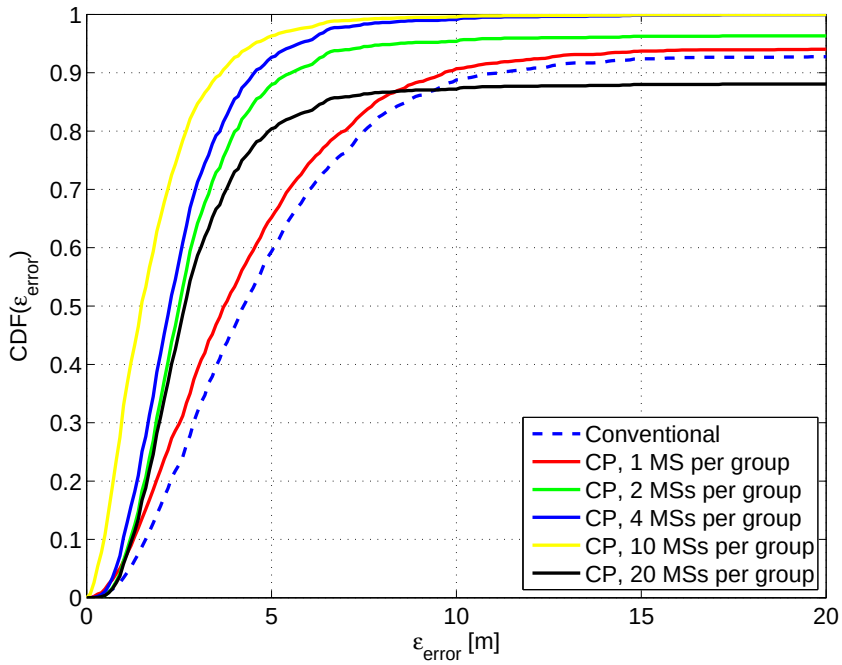


Figure 4.14: CP using EKF with realistic communications constraints and different numbers of MS per group.

too few ANs are available. On the other hand, if the number of ANs is too high (e.g., 70), the coverage by the ANs limits additional cooperation gains. In this case where the AN coverage is so good the cooperation gain is marginal.

In chapters 5, 6, and 7 the localization error is represented as a 2-dimensional symmetric Gaussian random variable, and it is interesting to see how well this assumption fits with the present results. The plot in Figure 4.15 therefore shows two CDF-curves of a 2-dimensional Gaussian position error. The error distance is calculated as: $\varepsilon_{\text{error}} = \sqrt{X^2 + Y^2}$, where X and Y are two 1-dimensional zero-mean Gaussian random variables. The Gaussian CDF curves have been hand-fitted to the results of the cooperative EKF algorithm, for 30 and 70 ANs. The curve for 30 ANs, shows quite good resemblance with the Gaussian error curve, however the curve for 70 ANs does not fit as well. Nevertheless, we assume that the value for $\text{CDF}(\varepsilon_{\text{error}}) = 0.68$ is a good approximation for the standard deviation, since for a Gaussian distribution this would correspond to the standard deviation. This result shows that the localization error standard deviation can be as little as 2-3 meters, with a sufficient deployment of ANs of between 30 and 70 ANs per 100x100 square meters. In

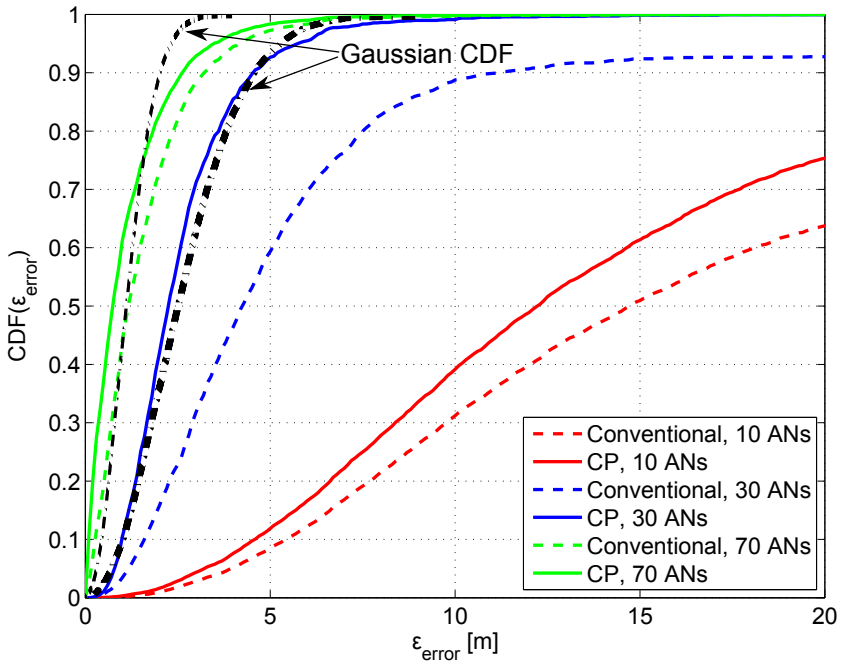


Figure 4.15: CP using EKF with realistic communications constraints and different numbers of ANs.

comparison to the results in Figure 2.5 on page 27, where an estimated standard deviation for a GPS systems approximately 13 meters and the result for the hybrid localization scheme that is using GPS+Galileo+Cellular is 8 meters, the considered cooperative EKF scheme with UWB ranging measurements is clearly more accurate.

For further comparison, it is worth mentioning the survey of different localization systems in [Liu et al., 2007]. Here the mentioned results for Wi-Fi based systems are in the order of 1-3 meters for static positioning, however for tracking a walking user the only mentioned result is 5 meters with 90% probability. Further, they describe that Bluetooth based localization can be as accurate as 2 meters with 90% probability, however only for static localization.

4.7 Conclusion

In this contribution, we have analyzed cooperative positioning and tracking algorithms under realistic communications constraints. These constraints were

modeled here based on a Wi-Fi infrastructure and error models based on empirical measurements for UWB devices. It was shown that the introduction of realistic communications constraints resulted in an added delay, which had a significant effect on the positioning performance, especially for the cooperative algorithms. This is mainly due to the more complex measurement exchange that is necessary to realize the centralized cooperative positioning algorithms. We found that the static solution and the EKF algorithms were similarly affected by the realistic communications constraints. Further, we observed that increasing the number of cooperating MSs had a positive impact on the positioning performance, as expected due to added cooperation possibilities. However, this was only until a tipping point was reached and the performance became worse with additional cooperating MSs. This tipping point is likely a result of the communication overhead becoming large, which in turn leads to increased delays. Nevertheless, in most cases the cooperative approach strongly outperforms the conventional (non-cooperative) approach. Specifically, the evaluation results showed that the localization error was quite close to being Gaussian distributed and that a standard deviation of the tracking error of 2-3 meters represents the performance of the cooperative EKF algorithm, if a sufficient number of anchor nodes is deployed in the considered scenario.

References

F. W. C. Chan and H. C. So. Accurate distributed range-based positioning algorithm for wireless sensor networks. *IEEE Transactions on Signal Processing*, 57(10):4100–4105, October 2009.

Q. Chen, F. Schmidt-Eisenlohr, D. Jiang, M. Torrent-Moreno, L. Delgrossi, and H. Hartenstein. Overhaul of IEEE 802.11 modeling and simulation in ns-2. *Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems*, 2007.

S. Frattasi. *Link Layer Techniques Enabling Cooperation in Fourth Generation Wireless Networks*. PhD thesis, Aalborg University, Aalborg, Denmark, September 2007.

S. Gezici, Z. Tian, G.B. Giannakis, H. Kobayashi, A.F. Molisch, H.V. Poor, and Z. Sahinoglu. Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks. *Signal Processing Magazine, IEEE*, 22(4):70–84, 2005.

H. Liu, H. Darabi, P. Banerjee, and J. Liu. Survey of wireless indoor positioning techniques and systems. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(6):1067–1080, 2007.

C. L. F. Mayorga, F. della Rosa, S. A. Wardana, G. Simone, M. C. N. Raynal, J. Figueiras, and S. Frattasi. Cooperative positioning techniques for mobile localization in 4g cellular networks. *Proceedings of the IEEE International Conference on Pervasive Services*, July 2007.

C. Mensing and J.J. Nielsen. Centralized cooperative positioning and tracking with realistic communications constraints. In *Positioning Navigation and Communication (WPNC), 2010 7th Workshop on*, pages 215–223. IEEE, 2010.

M. Pezzin and A. Alvarez Vazquez. Real life ranging measurements with low power, low data rate transceivers. *Proceedings of the Workshop on Positioning, Navigation and Communication (WPNC)*, March 2010.

Marios Raspopoulos, Stavros Stavrou, Bernard Uguen, Roxana Burghelea, Mariano García, Troels Pedersen, Gerhard Steinböck, Bernard H. Fleury, Benoît Denis, Joe Youssef, Thierry Tenoux, Turgut Mustafa Oktem, and Dirk Slock. WHERE D4.5 - Modelling of the Channel and its Variability. Technical report, ICT-217033 WHERE, 2010.

H. Wymeersch, J. Lien, and M. Z. Win. Cooperative localization in wireless networks. *Proceedings of the IEEE*, February 2009.

5

Location based Relay Selection

The contribution of this chapter targets the problem of mobile relay selection in the small scale scenario presented in Figure 1.2. Specifically, this contribution focuses on how inaccurate input information affects the performance of location based and SNR based relay selection schemes. The relaying schemes are used to enhance the throughput of downlink transmissions from the AP to mobile destination nodes, by using two-hop relaying in cases where the direct link performance is suboptimal due to bad link conditions. An SNR-measurement based scheme and a location information based scheme are proposed and compared in terms of how delayed and inaccurate state information and model parameters impact the performance of the schemes.

5.1 Introduction

In wireless networks the performance of data transmissions is depending on the distance between the transmitter and the receiver as well as on the number of collocated nodes and their activity patterns. In cases where such phenomena cause poor link quality between a transmitter and receiver pair it is well known in the literature that relaying techniques can be used to improve performance, as described in chapter 2, However, as shown in Figure 5.1, there may be many relay candidates to choose from.

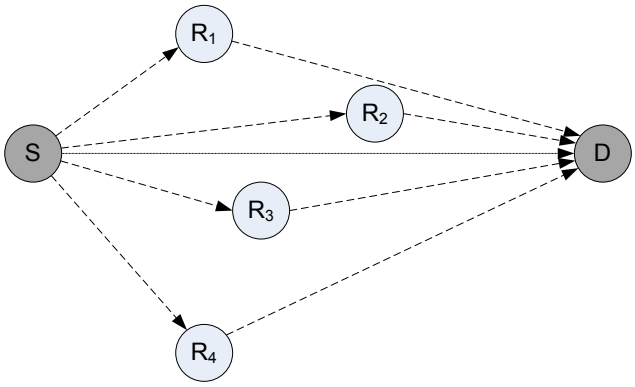


Figure 5.1: For any given source and destination pair of a relayed transmission, there may be multiple candidate relays to consider.

Choosing the best relay node in a given situation requires information that can be used to judge the quality of the concerned links, for example the links shown with dashed lines in Figure 5.1. Typically, this information is updated regularly, which allows to look up the most suitable relay node for a frame transmission from a table of historic measurements. In general, such a relaying system can be illustrated as in Figure 5.2.

As indicated in Figure 5.2, there are multiple error sources in such a relaying system. In addition to inaccuracies due to measurement error and variations, changes in the environment, such as node movements, can lead to the historic information becoming outdated and thus inaccurate. Figure 5.3 exemplifies how node movements can make historic measurements inaccurate.

The consequence of node movements has been investigated for the CoopMAC relaying protocol in [Liu et al., 2007] and it is found that as the average node movement speed increases, the performance of the scheme decreases due to outdated information on neighboring nodes' suitability as relays. One way of improving this situation would be to make sure that measurements are updated frequently enough to cancel out the effect of node movements, however since

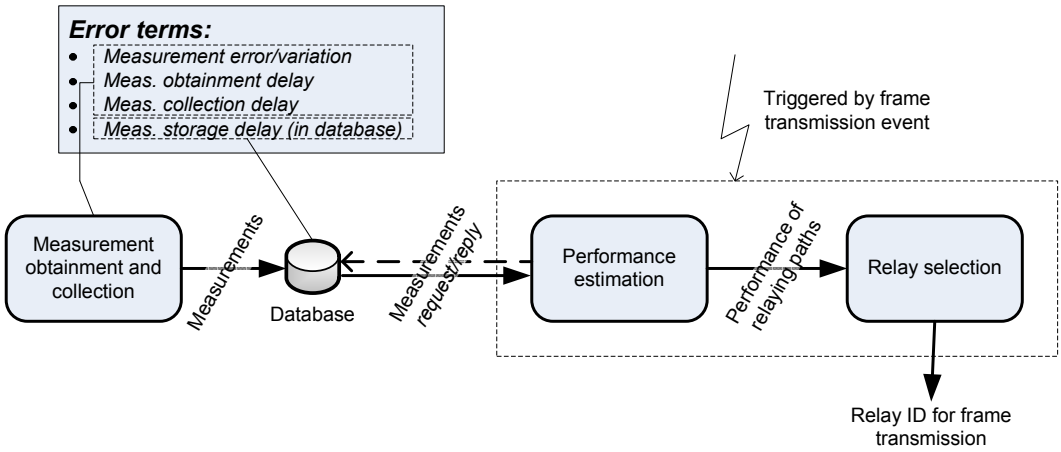


Figure 5.2: Generalized relaying system where the relay to use for a frame transmission is determined from historical measurements that may be affected by the listed error terms.

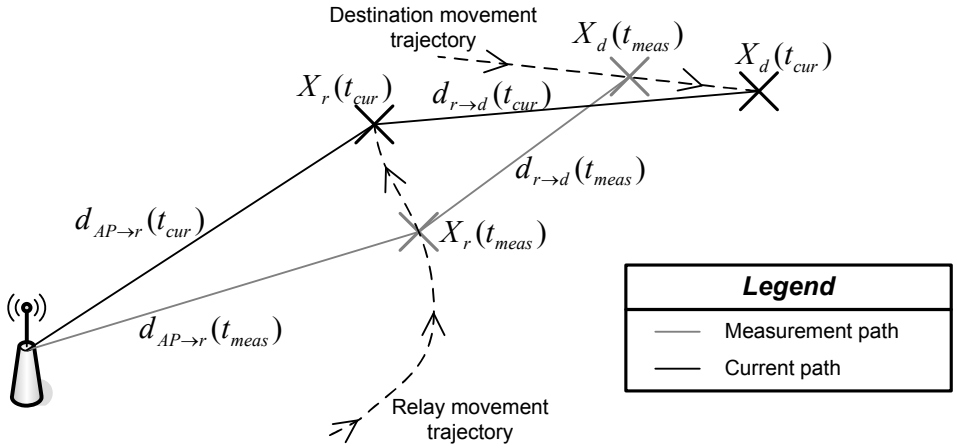


Figure 5.3: Node movements and aged measurements introduce errors in the node position and link distance knowledge. X denotes a node position, d is a link distance, t_{meas} is the time of the latest measurement, and t_{cur} is the current time.

this will increase the signaling overhead related to collecting measurements, it is not immediately clear how to best trade-off this setting for a relaying system. The CoopMAC protocol uses information on the achievable bit rate on each link from past frame transmissions for relay selection, but in principle other types of information could be used as well. Since location information is often available on personal devices such as smartphones or tablets, as mentioned in chapter 1, this type of information may be useful for relay selection. In this chapter a location-based scheme and a link-quality based scheme are therefore

compared and evaluated with respect to the impact of node movements. Further, since the information used for relay selection is not necessarily completely accurate, e.g., due to measurement error/variations, analyzing this aspect will also be a part of the analysis. Finally, since the use of location information requires a model that can estimate the performance based on node positions in order to perform relay selection, it is interesting to study how inaccurate model parameters impacts the relay selection.

5.2 Related Work

Within the field of relaying, much work has been done, first from an information theoretic point of view and more recently from a practical network point of view. In this section an overview of examples of existing work is given.

Initially, some examples of geographic relaying/forwarding protocols are given. The GeRaF protocol presented in [Zorzi and Rao, 2004] is functionally very similar to the Harbinger protocol described in [Zhao and Valenti, 2005]. Both protocols assume that nodes have perfect knowledge of their position and that the position of the final destination is contained in the packet header. The next hop node is however not known on beforehand. When potential relays overhear the transmission, they let their contention time depend on how close they are to the destination node, so that the receiving node closest to the destination will act as the relay and the packet is forwarded towards the destination. In practice, this type of protocols does not easily allow for rate adaptation since the transmission distance for each hop is unknown, and a transmitting node would not be able to judge which transmission mode is best on beforehand. Even if the packet is forwarded successfully using the highest bit rate, it could in some case be beneficial to use a lower bit rate, since the packet could travel farther for each hop and thus require fewer hops to reach the final destination.

An often referenced example of an IEEE 802.11 DCF based relaying scheme, is the rDCF scheme described in [Zhu and Cao, 2006]. The scheme is basically an extension of IEEE 802.11 DCF to support multi-hop transmissions. If a node overhears a transmission between two other nodes, and it knows that it can achieve higher bit rates towards each of these nodes from previous transmissions, it will add the sender and receiver to its willing list. The willing list is periodically updated. Upon transmission, the sender coordinates with the most suited relay, through a triangular handshake.

A quite similar example of existing work is the coopMAC protocol [Liu et al., 2007], which is also targeted at IEEE 802.11 based wireless ad hoc networks. The protocol monitors and stores the achievable bit rates for transmissions between peers in its cooperation table. The information in these tables are regularly exchanged between neighbor nodes. Upon transmission, a node looks up potential helper nodes in the table, and uses an extended but backwards compatible RTS/CTS handshake notify potential helper nodes. As shown in [Liu et al., 2007], mobility outdates information about potential helper nodes, which causes performance to decrease to the same level as if relaying was not considered. Even with low mobility (max 1 m/s, 60 s pauses), the gain compared to standard 802.11 without relaying is less than 10%. Compared to the rDCF protocol, CoopMAC is backwards compatible with IEEE 802.11 DCF, and can co-exist with such legacy nodes.

A further extension of rDCF is given in [Lu et al., 2009], where PRO (Protocol for Retransmitting Opportunistically) is proposed. Like CoopMAC and rDCF it monitors link quality of overheard packets, stores these observations, and periodically shares the information to the network. However, PRO uses the IEEE 802.11e EDCA protocol to let high quality relays retransmit with higher probability.

Finally, in [Tan et al., 2007] the protocol CODE is proposed. Also this protocol uses the same principle as rDCF and CoopMAC for collecting information about potential relays, but uses relaying to improve poor links using VAA (virtual antenna arrays. VAA increases the achievable bit rate through increased power gain. The fine synchronization that VAA requires, is achieved by a beacon based synchonization scheme. The protocol is also backward compatible with IEEE 802.11, and the paper shows throughput improvements of up to 60 %.

In [Liu and Wong, 2004] the Relay-Based Adaptive Auto Rate protocol (RAAR) is presented. This protocol targets the performance anomaly in 802.11, in which nodes located far from the AP pull down the performance of the whole network. By introducing forwarding relays for the nodes close to the edge of the coverage area, the throughput of the whole network goes up. Only upstream traffic in the Wide-Area Network (WAN) is considered in this paper. The authors mention that the relay selection can be done based on location information and current modulation scheme, which is exchanged through periodical status reports. They also mentions that the data may be outdated due to mobility or channel conditions.

A different approach is taken with the CCMAC protocol presented in [Hu and Tham, 2010], which aims at improving throughput for uplink transmissions (client to AP) in the region near the AP by allowing simultaneous source to relay transmissions. With this protocol, the AP is responsible for coordination and enabling of concurrent transmissions. The applied algorithm uses reinforcement learning, which means that performance increases over time, as the algorithm learns which relay configurations are best.

In [Narayanan and Panwar, 2007] the authors analyze two-hop relaying with regard to throughput, delay, and energy consumption. The authors conclude that two-hop relaying is not more energy demanding than direct transmission, even though some nodes need to act as relays and perform extra relaying transmissions, if the cost vs. benefit is considered on a large time scale.

A general problem for the rDCF and CoopMAC like protocols, is that they are intended for mostly stationary scenarios. If nodes are moving, the entries in the cooperation tables become deprecated, as shown clearly in [Liu et al., 2007]. Another option is to use location information as mentioned in [Liu and Wong, 2004], however the authors also give a word of caution since this information can be outdated due to, e.g., mobility. The RAAR and CCMAC protocols presented in [Liu and Wong, 2004, Hu and Tham, 2010] both focus on improving the uplink throughput. However, in most cases where people are downloading or streaming content from the internet, the majority of data is moving in the downlink direction from the AP to the user node. In the contribution presented in this chapter, relaying is considered for improving throughput in the downlink direction and both link measurements and location information is considered for the relay selection.

5.3 Centralized Relay Selection

In this work we consider the scenario where downlink transmissions are made towards mobile users, as sketched in Figure 5.4. Audio or video streaming are examples of applications leading to such traffic patterns. In such downlink scenarios we will therefore focus on *centralized two-hop relay selection* where it is assumed that the AP is the source of all transmissions. Data transmissions can be done directly to the destination MD or as a two-hop transmission via any intermediate relay node, as sketched in Figure 5.4. Whether the transmission is direct or relayed, is decided by the relaying schemes.

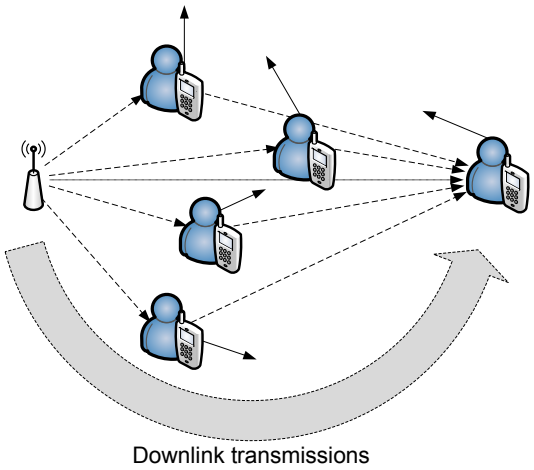


Figure 5.4: Example of downlink scenario where transmissions are made from AP to a mobile destination node, either as a direct transmission or via a mobile relay.

Generally, the relaying scheme should decide to make a transmission via a relay if this path provides a better transmission quality than a direct transmission. Formally, if Q denotes the transmission quality, this can be expressed as:

$$Q_{\mathrm{AP}\to D} < Q_{\mathrm{AP}\to\mathrm{R}\to\mathrm{D}} \tag{5.1}$$

Further, in order to achieve the best performance, the relay R_{opt} that provides the highest quality transmission should be used:

$$R_{opt} = \underset{R}{\operatorname{argmax}}\, Q_{\mathrm{AP}\to\mathrm{R}\to\mathrm{D}} \tag{5.2}$$

The quality of a transmission can be quantified using the most appropriate metric for the application, for example low BER, low delay, or high throughput. In this work two metrics are considered for the relaying condition in equation (5.1). One is based on the BER, where the BER of the relayed path is calculated as:

$$BER_{\mathrm{rel}} = 1 - (1 - BER_{\mathrm{AP}\to\mathrm{R}}) \cdot (1 - BER_{\mathrm{R}\to\mathrm{D}}) \tag{5.3}$$

where the BER of each link is estimated from the SNR, using theoretical expressions for the specific modulation scheme and Ricean channel from for example [Proakis, 1995]. For efficiency, it would also be possible to pre-generate a look-up table with expected BER for different SNR levels and Ricean K-factors.

The second selection metric is based on the expected throughput, which is calculated using the throughput model described in chapter 3. The throughput of a direct transmission is calculated as:

$$S_{\text{dir}} = \frac{P_{\text{suc}}^{\text{AP}\to\text{D}} \cdot B_{\text{MSDU}}}{E[T_{\text{tx}}^{\text{AP}\to\text{D}}]} \tag{5.4}$$

and the throughput of the relayed path is calculated as:

$$S_{\text{rel}} = \frac{P_{\text{suc}}^{\text{AP}\to\text{R}} \cdot P_{\text{suc}}^{\text{R}\to\text{D}} \cdot B_{\text{MSDU}}}{E[T_{\text{tx}}^{\text{AP}\to\text{R}}] + E[T_{\text{tx}}^{\text{R}\to\text{D}}]} \tag{5.5}$$

where $P_{\text{suc}}^{\text{x}\to\text{y}}$ is the probability of a successful frame transmission from node x to node y, $E[T_{\text{tx}}^{\text{x}\to\text{y}}]$ is the expected time of a frame transmission from node x to node y, and B_{MSDU} is the frame payload size.

In order to evaluate the quality condition above, it is necessary to collect measurements of the quality of the different paths. In this chapter, two types of measurements are considered for this purpose. The first is a measurement of the link quality between the mobile nodes in the form of SNR measurements, whereas the second type constitutes location measurements of the mobile nodes. In the following, two relay selection schemes are considered that correspond to the two types of measurements.

5.3.1 SNR based Relay Selection

In order to obtain a view of the link quality of all possible relay paths, the AP needs to collect SNR measurements for all links involved in the possible relay paths. This is done by using the following two-step approach for each mobile node, which is sketched in Figure 5.5. First, the mobile node broadcasts a *hello* message, which allows the neighboring nodes to measure the RSS and hereby estimate the SNR on the link towards that node. It is assumed that all nodes are using the same fixed transmission power for the hello broadcasts. This measurement is assumed to represent the link state at this moment in time. Secondly, the neighboring nodes that overhear the hello broadcast send *measurement* messages that contain the observed SNR measurements to the AP. These hello broadcasts are generated by every mobile node periodically with cycle time μ_{hello}[s].

Hereby N nodes that each send out a hello broadcast result in $N \cdot (N-1)$ measurement transmissions to the AP, in the case where all nodes receive all hello broadcasts. The actual amount of measurement transmissions may vary due to losses and possible retransmissions. As this approach generates many

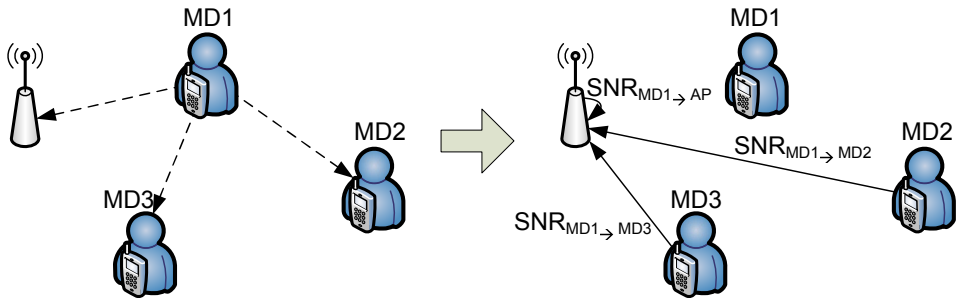


Figure 5.5: Based on a mobile device's broadcasted hello message, the neighboring nodes are able to measure the link SNR, which they transmit back to the AP.

individual transmissions that contribute to the overhead, future work could consider to accumulate a bulk of measurements at each MD before initiating a transmission to the AP.

Whenever the AP has a frame transmission scheduled, it uses the available collected measurements to determine the if the transmission should be direct or relayed and if so, which relay to use. This is done using the equations (5.1) to (5.5) given above, after calculating the expected BER or expected throughput from the collected SNR measurements as described in chapter 3. However, since old measurements may be misleading the AP has an age limit on the storage time of measurements specified by the α_{store} parameter. Measurements older than this threshold are discarded and not used for relay selection.

5.3.2 Location based Relay Selection

The idea behind this scheme is that by knowing the locations of the MDs in the network, the path-loss, SNR and in turn the BER can be estimated with propagation models by assuming fixed transmit power and approximating propagation properties of the environment. Locations are obtained by letting all MDs transmit location measurements periodically with interval μ_{loc} to the AP using unicast transmissions, as sketched in Figure 5.6.

Whenever the AP has a frame transmission scheduled, the approach is basically similar to the SNR based relay selection in the previous section 5.3.1. Also location information is discarded if too old. However, the collected location information needs to be converted into link BER or link throughput estimates to be useful for relay selection. First the path-loss is estimated using the path-loss model presented in section 2.1.1. As the value of the path loss

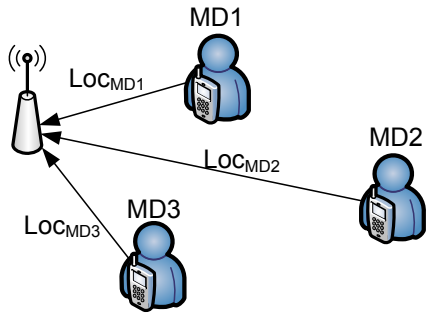


Figure 5.6: Location measurements are collected by sending them from each mobile device to the AP.

exponent n is scenario dependent and its exact value is typically not known in advance, we will investigate the sensitivity to inaccurate estimates of this parameter in section 5.6.2.

Given a specific transmit power level P_{tx}, the calculated path-loss $\overline{PL}(d)$ and assumed noise floor N_{floor}, the SNR is calculated as:

$$SNR = P_{\text{tx}} + \overline{PL}(d) - N_{\text{floor}} - X[dB] \tag{5.6}$$

where X is a random variable representing shadowing due to obstacles in the environment. Initially we will assume $X = 0$, however in NLOS situations that we will investigate later in this section it will be necessary to guess the attenuation. This is covered in section 5.6.2. Having determined the SNR, the expected BER or expected throughput can now be calculated as described in chapter 3, and used for relay selection as described in the equations (5.1) to (5.5).

5.4 IEEE 802.11 based Evaluation Scenario

In the following, the presented relay selection schemes are evaluated in a specific IEEE 802.11 based scenario, since the 802.11 technology is the de facto standard for local wireless networks; typically denoted Wi-Fi. However, the 802.11 technology does currently not support relaying when working in infrastructure mode, so the proposed relaying schemes are foreseen as MAC protocol extensions as described in the following.

In this scenario only a single AP and its associated mobile Wi-Fi devices are considered, and it is assumed that there is no inter-cell interference from neighboring cells. Since IEEE 802.11a supports 11 independent channels, this

assumption is considered reasonable in situations where channels have been allocated properly.

5.4.1 SNR Measurement Collection

The hello message is a IEEE 802.11 MAC frame without payload, since only the MAC address is needed for the receiver to identify the broadcast source. The frame has a total size of 20 octets. Notice that hello broadcasts may be lost if collisions occur. In this case the state of the corresponding link is not updated and the last successfully received measurement is used. We can further define the age of a measurement as the elapsed time since the hello broadcast of the latest measurement for that link was initiated. The age of a link measurement is a stochastic process that is influenced mainly by the hello broadcast generating process. For this evaluation, a random jitter is added to the inter-event time for hello broadcasts, which ensures that hello transmissions from different MDs are not in sync.

As described in section 5.3.1, whenever an MD overhears a hello broadcast and thereby obtains an SNR measurement, it assembles a measurement frame and sends it to the AP using a unicast transmission. Due to the small frame size, RTS/CTS is not applied but standard 802.11 MAC retransmissions are used if needed. The measurement frame is envisioned as being a MAC control frame that carries the MAC address of the hello broadcast source (6 octets) and the SNR measurement (2 octets), which amounts to a frame size of 28 octets when adding this information to the standard 802.11 control frame layout [IEEE, 2007].

The AP identifies the link from which the measurement has been obtained from the MAC addresses of the broadcast node and measurement node. Notice that it is assumed that links are symmetric, meaning that for example the link between MD1 and MD2 is updated both when MD2 overhears a broadcast from MD1 and when MD1 overhears a broadcast from MD2.

Having obtained SNR measurements from links between devices in the network, the AP can now chooses the best path.

5.4.2 Location Measurement Collection

Similarly to the case with hello broadcasts for the SNR measurement based scheme, the initial transmission time is chosen for each node uniformly random

in the interval $[0, \mu_{\text{loc}}]$. Further, the following location measurement transmissions are offset with a uniform random jitter in the interval $[-0.1 \cdot \mu_{\text{loc}}, 0.1 \cdot \mu_{\text{loc}}]$ to avoid transmissions being in sync. The measurement frame is a MAC control frame that carries the longitude (4 octets) and latitude (4 octets) of the node. Assuming the longitude and latitude are given as a degree decimal fraction and the circumference of the earth is $40000km$, the precision that is supported by this format is approximately $\frac{40000km}{2^{4 \cdot 8}} = 0.01m$, which is judged to be sufficiently precise. The frame size amounts to 28 octets when adding longitude and latitude information to the standard 802.11 control frame layout [IEEE, 2007]. This is the same size as the SNR measurement frame.

5.5 Evaluation Methodology

For evaluation we consider simulations of mobility and the wireless network followed by a combined performance evaluation as sketched in Figure 5.7.

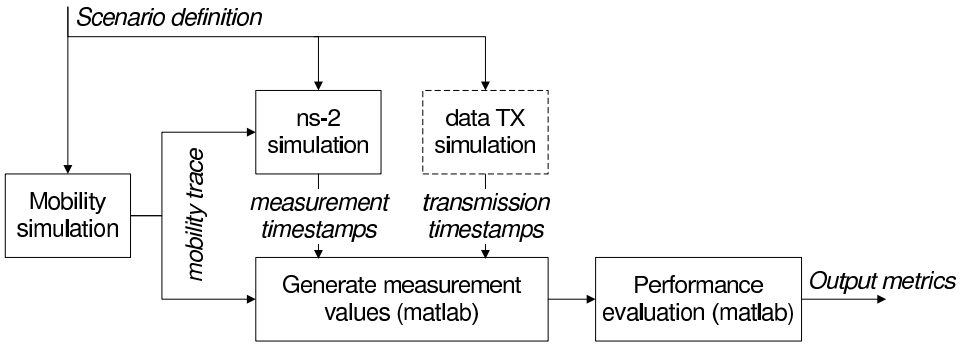


Figure 5.7: Simulation overview.

First a simulation of node mobility is generated based on the random waypoint mobility model, described in chapter 2. The outcome is a trace of the movements of all nodes.

5.5.1 ns-2 Simulation

Now the ns-2 simulation[1] is executed, based on the mobility trace and the scenario specific parameters listed in Table 5.1. We use the *802.11ext* module

[1]The ns-2 simulation is based on [Chen et al., 2007], which has been updated with the author's patch from October 21, 2008.

to simulate realistic 802.11a behavior. This ns-2 version includes a Nakagami fading model which has been parametrized according to Table 5.1 with model parameters $\Gamma = n$ and $m = \frac{(K+1)^2}{2K+1}$ to approximate a Ricean fading environment.

For evaluating the SNR and location based algorithms, we use two different custom ns-2 agents for generating and collecting measurements. For the SNR based algorithm the agent in each MD periodically generate broadcasts and forward overheard broadcasts to the AP, as described in section 5.3.1. For the location-based algorithm, the agent makes each MD transmit a measurement frame containing its location periodically as described in section 5.3.2. We assume that MDs are able to obtain their own location coordinate $(x, y) + \epsilon_{\text{pos}}$ where ϵ_{pos} is a zero mean symmetric two-dimensional Gaussian error with standard deviation σ_{pos} representing the localization inaccuracy.

The outcome of the ns-2 simulation is a trace file that for every node describes when hello and measurement frames are received.

Scenario size	$100m\ x\ 100m$
No. of mobile devices	10
Mobility model	Random Waypoint (speed: $2 - 8m/s$)
Rice K-value	6
Path loss exponent n	2.9 (based on [Durgin et al., 1998], outdoor meas.)
Modulation scheme	BPSK (6 Mbit/s in 802.11a)
Noise floor	-86 dBm
Transmission power	100 mW
σ_{pos}	$0m$

Table 5.1: Scenario parameters.

5.5.2 Data Transmission Simulation

In this work the measurement collection and data transmissions are simulated separately in ns-2 and matlab, to make the implementation simpler and thus allow for rapid prototyping. The considered solution does therefore not take into account the mutual influence of the data transmissions and the measurement collection and a future work item is to take this interaction into account. We believe that even though this gives a slightly optimistic view of the achieved performance, the analysis will still give valuable insights into the benefits and drawbacks of measurement based and location based relaying algorithms. With this simplification, this step simply constitutes the creation of

a list of timestamps and corresponding destination nodes. The timestamp and destination pairs in the list, specify when and to which mobile node the data transmissions are done. The list of data transmissions is created with an exponentially distributed inter-transmission time, with the rate being: $\frac{1}{\mu_{\text{tx}}}\ s^{-1}$. For each transmission, a destination node is selected randomly among the mobile nodes. The μ_{tx} inter-transmission time means that there is mostly some idle time between transmissions. The performance results presented later in this chapter, are based on an average of the BER or throughput achieved for the transmissions. The idle time between transmissions is not included in these results.

5.5.3 Generate Measurement Values

The purpose of this step is to generate the measurement values that correspond to the timestamps of measurement availability, given by the output from the ns-2 simulation. Here, SNR and location measurements are generated, taking into account the influence of node movements and localization inaccuracy. The measurements are generated as described in the following. For each data transmission, the generated mobility model trace is processed, to determine exactly where the mobile nodes are located at the time instant where the data transmission is initiated. These positions are the true positions of the nodes at the transmission time t_{tx}, and it is denoted as $X_n(t_{\text{tx}})$, where n is the node id.

As the measurement based schemes are using collected measurements, it is necessary to generate the measurements so that they correspond to the system state for when the measurements were actually obtained, t_{meas}.

Location Measurements

Now, the location measurement for node n at transmission time t_{tx} is given as:

$$\hat{X}_n^{\text{loc}}(t_{\text{tx}}) = X_n(t_{\text{meas,loc}}) + \epsilon_{\text{pos}} \tag{5.7}$$

where $X_n(t_{\text{meas,loc}})$ is the true node position at the time the measurement was collected $t_{\text{meas,loc}}$, and ϵ_{pos} is a zero mean symmetric two-dimensional Gaussian error with standard deviation σ_{pos} representing the localization inaccuracy. This assumption is in line with the findings in Chapter 4 where the location error was found to be approximately Gaussian distributed. The true node position at the time a measurement was collected, $X_n(t_{\text{meas,loc}})$, is calculated from the generated mobility trace.

SNR Measurements

Like for the location based scheme the true node position for when a measurement was obtained $X_n(t_{\text{meas,loc}})$, is used. However, since SNR measurements relate to a link between two nodes, the positions of both nodes (n_1 and n_2) are needed. The SNR measurement is generated by first determining the distance between the nodes as:

$$d_{(n_1,n_2)} = ||X_{n_1}(t_{\text{meas,SNR}}) - X_{n_2}(t_{\text{meas,SNR}})|| \tag{5.8}$$

and secondly to calculate the path loss and SNR using the path loss model presented in chapter 2 on page 19 and the SNR calculation in section 3.3 on page 45.

Notice that $t_{\text{meas,loc}}$ and $t_{\text{meas,SNR}}$ depend on the respective update frequencies, used for each of the schemes.

5.5.4 Performance Evaluation

In order to evaluate the performance of the SNR measurement based path selection scheme and the location measurement based path selection scheme, we compare the performance of these schemes to the following reference schemes: the case where the *direct* path is always used and the *ideal* case where exact and updated link state information is always available.

The first step in the performance evaluation is that the relaying decisions for each data transmission are calculated for each of the considered schemes, given the available measurements. The relaying decisions are determined from the conditions described in section 5.3 on page 80.

The second step is to apply the relaying decisions of each scheme on the actual link conditions for each data transmission, and thereby obtain the achieved performance for each scheme. The actual link conditions are calculated in essentially the same way as the measurements are generated and the relay decisions are made, but for the true node positions and without location error. First the actual link SNR is determined from the path loss between the nodes' true positions. The actual SNR values for all link are then used to evaluate the performance both in terms of average BER and average throughput:

The achieved BER of a link is calculated from the link SNR using the `berfading` function in the matlab communications toolbox assuming BPSK

modulation scheme and the Ricean fading model (K=6). For relayed transmissions, the path BER is calculated as specified in equation (5.3).

The achieved throughput is calculated from the link SNR using the throughput model, with the performance of the direct and relayed transmission being specified in equation (5.4) and (5.5), respectively. Also here BPSK modulation scheme and the Ricean fading model (K=6) is assumed, as well as 6 Mbit/s bit rate.

Another way to evaluate the considered relaying schemes, would be to implement the relay decision protocol in a network simulator such as ns-2 or OMNET++. This would allow to evaluate how well the relaying protocols perform in situations where both uplink and downlink transmissions are ongoing and nodes are competing for channel access. This would mean that collisions could occur, e.g., due to hidden and exposed node situations and this type of evaluation could give a more realistic view of the achievable throughput improvements. However, as the purpose of this work is not to propose relaying schemes that work better than the already existing schemes described in section 5.2, but rather to analyze the difference in impact of mobility, delay and inaccurate input information on the SNR based and location based schemes, it was not found to be necessary with a full-blown network simulation.

Besides BER and throughput also the signaling overhead is calculated for the considered relaying schemes, as described in the following section.

5.5.5 Signaling Channel Utilization

This metric gives the overhead spent on obtaining and collecting measurements as a fraction of channel capacity, assuming all nodes are in a single collision domain. First, the transmission time T_{data} of hello and measurement frames is calculated as a function of the number of bytes MAC payload B_{MSDU}, as specified in Appendix A. Here $B_{\text{MPDU}} = 20$ bytes for hello broadcasts and $B_{\text{MPDU}} = 28$ bytes for SNR and location measurements.

Now the signaling channel utilization is estimated as in equation (5.9) and equation (5.10), where N is the number of MDs.

$$U_{\text{snr}} = \frac{N \cdot (T_{\text{data}}(B_{\text{MPDU}} = 20) + (N-1) \cdot T_{\text{data}}(B_{\text{MPDU}} = 28))}{\mu_{\text{hello}}} \tag{5.9}$$

$$U_{\text{loc}} = \frac{N \cdot T_{\text{data}}(B_{\text{MPDU}} = 28)}{\mu_{\text{loc}}} \tag{5.10}$$

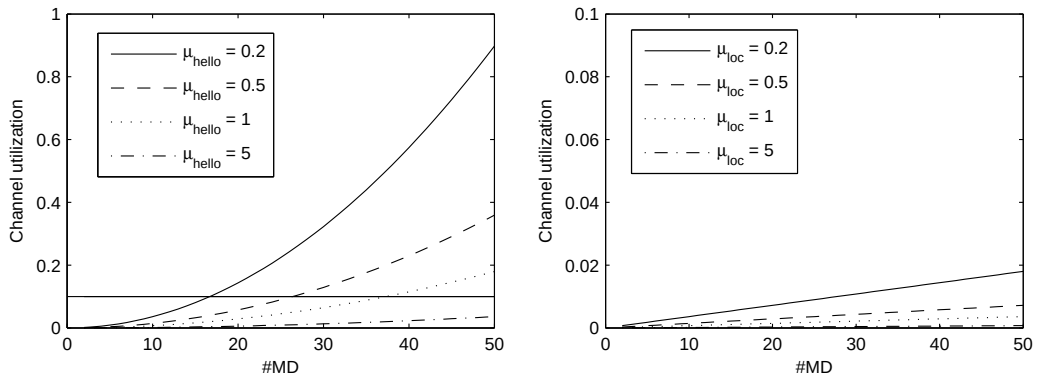
The actual signaling channel utilization may vary slightly due to possible collisions and retransmissions that are not accounted for in these calculations.

5.6 Results and Discussion

5.6.1 Signaling Overhead

Initially, we compare the overhead in terms of the channel utilization used for signaling for the SNR-based and location-based relaying schemes.

In Figure 5.8(a) the channel utilization spent for obtaining and collecting SNR measurements for different hello broadcast intervals and node densities is shown. This has been calculated by inserting different values of N in equation (5.9). As capacity should be used for data transmission and not spent as overhead for measurement collection, a utilization of more than 10% is considered unacceptable. In the plot we see that this limit is exceeded at slightly less than 40 MDs with $\mu_{\text{hello}} = 1s$. In vehicular scenarios where even faster updates are needed, we see that for $\mu_{\text{hello}} = 0.5s$ and $\mu_{\text{hello}} = 0.2s$ the utilization exceeds the 10% limit for just 25 and 18 MDs, respectively. This result emphasizes the need for a more efficient relay path selection scheme.



(a) SNR measurements

(b) Location measurements

Figure 5.8: Calculated utilization of channel for obtaining and collecting measurements (Data rate is 6Mbit/s). Notice that the y-axis in Figure 5.8(b) is only a tenth of the y-axis in Figure 5.8(a).

Turning our attention to the proposed location-based scheme, we see from the channel utilization plots shown in Figure 5.8(b) that the used overhead is much lower for this scheme compared the the SNR-based scheme. The

curves are calculated using equation (5.10). Here, the 2% utilization is never exceeded in the plot, even when we have 50 MDs with measurement intervals of just $\mu_{\text{hello}} = 0.2s$, which lead to a utilization of 90% for the SNR measurement based scheme. This is mainly due to the fact that the utilization of the location based scheme grows linearly with the number of MDs whereas the growth is quadratic for the SNR based scheme.

5.6.2 Performance of the SNR based Algorithm

The results have been created using the parameters and settings listed in Table 5.1 and Table 5.2. The default parameters for the ns-2 802.11ext model have been used if not explicitly specified in the tables. The errorbar in the results show the overall mean and 95% confidence intervals for the mean values obtained in each simulation run.

Simulation time	$360s$
No. of simulation runs	16 - 32
Hello interval μ_{hello}	$5 \pm \text{uniform}(0..0.5)s$
Location interval μ_{loc}	$f(N, \mu_{\text{hello}})$, see section 5.6.3
Transmission interval μ_{tx}	0.5 s (exponentially distributed)
Storage time α_{store}	20 s

Table 5.2: Simulation parameters in addition to Table 5.1.

Varying number of MDs

The first results in Figure 5.9 show the achieved BER for varying number of MDs. We see that increasing the number of MDs does not have a practical impact on the relative performance of the ideal and measurement based schemes. However compared to the always direct scheme, the measurement based and ideal schemes are gaining better BER performance. This demonstrates that when the node density increases, further relay transmissions via short links are possible, which in turn leads to reduced BER.

Varying hello interval

The next set of results shown in Figure 5.10 show the impact on the measurement based scheme of varying the hello interval μ_{hello} compared to the ideal and direct schemes.

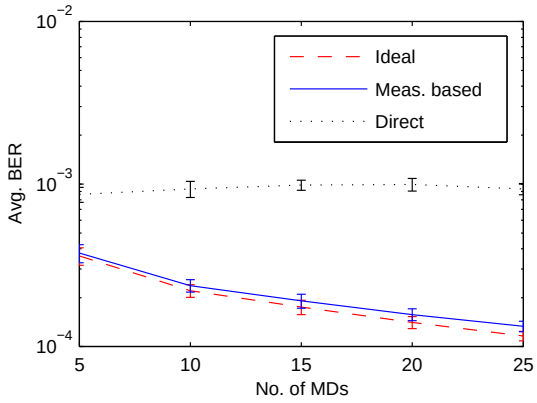


Figure 5.9: Achieved avg. BER for varying no. of MDs.

In Figure 5.10 we see that for $\mu_{\text{hello}} < 5s$, the achieved BER of the measurement based scheme is very close to the ideal scheme. As μ_{hello} increases, the measurement based scheme tends towards the direct scheme. It is also noteworthy that the BER never seems to increase above the level of the direct scheme. This can be explained by the fact that all measurements that are older than a predefined storage threshold are deleted and thus, they do not lead to an uncorrect decision. By applying such storage threshold the entries with the stale information are effectively discarded.

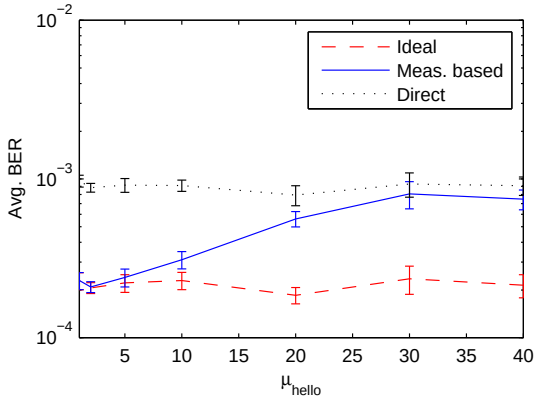


Figure 5.10: Achieved avg. BER for varying μ_{hello}.

Varying node speed

We now analyze how the mobility model impacts the path selection. Figure 5.11 shows the achieved BER when the speed of the mobile devices is increased.

The speed is varied from pedestrian up to vehicular speeds. The x-axes in both figures represent the average speed according to Table 5.3. Figure 5.11 clearly

	Speed [m/s]		
ID	min	max	avg.
1	0.5	2	1.25
2	2	8	5
3	5	15	10
4	10	20	15

Table 5.3: Minimum and maximum speed used for the RWP mobility model and the corresponding ID and average speed.

shows that increasing the mobility speed leads to a significantly worse BER peformance than the direct scheme. This is of course highly undesirable, as the AP would be better off by using only direct transmissions. One obvious solution would be to increase the hello broadcast rate, as this would reduce the age of measurements. A downside to this is that the signaling overhead increases linearly with the hello broadcast rate. Increasing the broadcast rate is therefore costly in terms of capacity. We therefore investigate if limiting the storage time with the α_{store} parameter can improve this situation without increasing the signaling overhead.

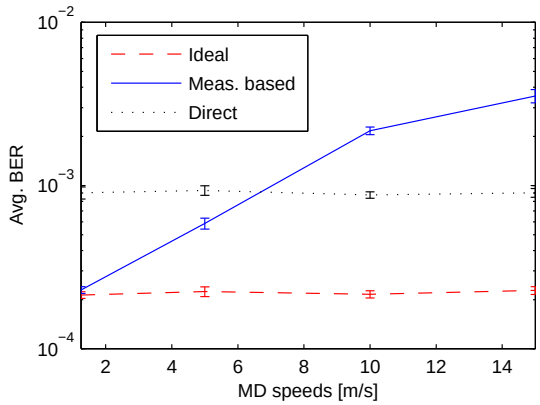


Figure 5.11: Achieved avg. BER for varying mobility speed.

Varying storage threshold

The results in Figure 5.12 show the achieved avg. BER for the mobility model with parameters specified by ID 3 in Table 5.3, which has a significantly worse

performance than the direct scheme in Figure 5.11. In Figure 5.12 we see that setting $\alpha_{\text{store}} = 2s$ the resulting avg. BER is slightly lower than the direct scheme. Hereby we have shown that by proper setting of the α_{store} parameter it is possible to enhance performance without additional signaling overhead for a given scenario. As $\mu_{\text{hello}} = 5s$ in the considered scenario, setting $\alpha_{\text{store}} = 2s$ entails that the AP does not always have knowledge of a link, but measurements are only used when they are fresh, i.e. less than 2 seconds old. Further, as we consider symmetric links, both measurements from the two MDs of a link contribute to the AP's knowledge of a link. this means that the fraction of time where the AP has knowledge of a specific link, lies in the interval between $\frac{2}{5}$ and $\frac{4}{5}$, since hello broadcasts are unsyncronized and assumed to be i.i.d.

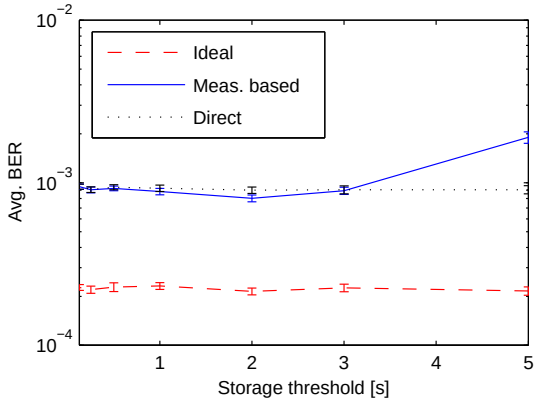


Figure 5.12: Avg. BER for varying α_{store}, for mobility model ID 3.

5.6.3 Performance of the Location based Algorithm

In this section now also the location-based relay selection algorithm is considered and compared to the SNR based scheme. Both results for the BER and throughput relay selection schemes and for the BER and throughput performance metrics are shown, in order to understand more clearly the difference between the selection schemes.

Varying node speed

Figure 5.13 shows that for the same measurement rate ($\mu_{\text{hello}} = \mu_{\text{loc}}$) the SNR and location based schemes perform similarly. However, if we instead use the μ_{loc} that satisfies $U_{\text{snr}} = U_{\text{loc}}$ (see equation (5.9) and equation (5.10)) the

signaling overhead in terms of channel utilization will be the same for the SNR and location based schemes. The performance of the location based scheme will in this case be close to the ideal scheme.

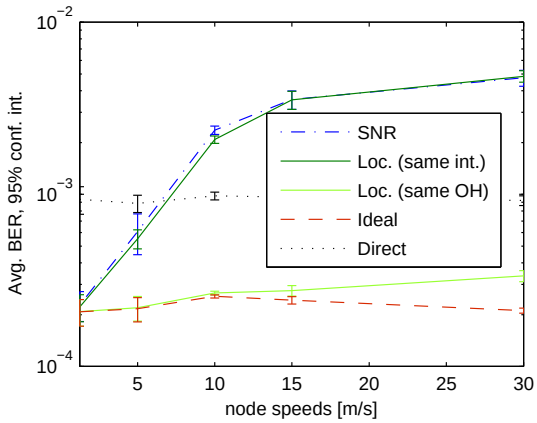


Figure 5.13: BER impact of varying node speed.

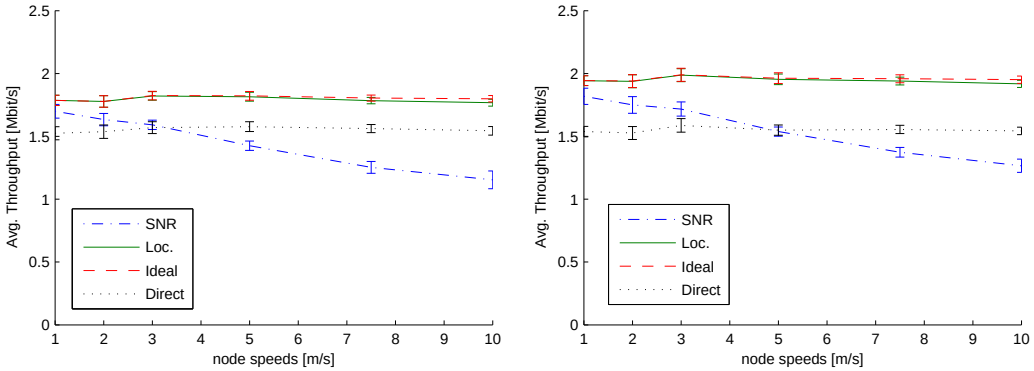
For all following results, the location based scheme uses the μ_{loc} that satisfies $U_{\text{snr}} = U_{\text{loc}}$.

Considering the same scenario, though with different mobility speed, and the BER based selection algorithm, the achieved performance expressed in terms of throughput is shown in Figure 5.14(a). If we compare this to the corresponding result in Figure 5.14(b) where the relay selection is based on throughput, we can notice some differences between the results. Generally, the throughput based selection leads to a slightly higher throughput for the ideal and location based algorithms, as well as for the SNR based algorithm. The direct scheme is, as expected, unaffected by the choice of selection criteria.

Notice that in the following results, a constant movement speed of 5 m/s is used.

Varying location accuracy

As the accuracy of location information estimates depends on uncontrollable factors and therefore may vary, we investigate the achieved performance for different accuracy levels in Figure 5.15. The path loss model used for the location based scheme uses the true path loss exponent $n = 2.9$ in this plot. Figure 5.15 shows that the location measurement based scheme performs close to the ideal scheme for $\sigma_{\text{pos}} < 5m$, while it is still better than the SNR based scheme up to $\sigma_{\text{pos}} < 10m$ and becomes worse than the direct scheme when



(a) BER based relay selection (b) Throughput based relay selection

Figure 5.14: Throughput impact of varying node speed.

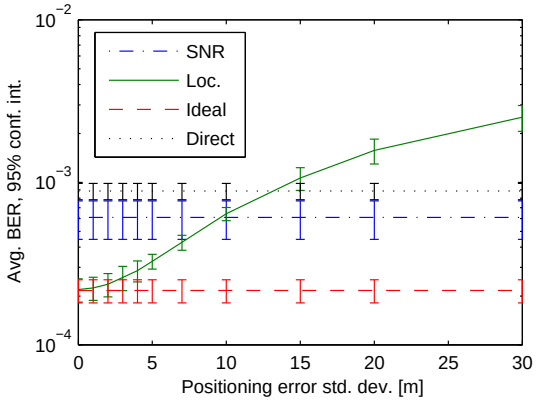
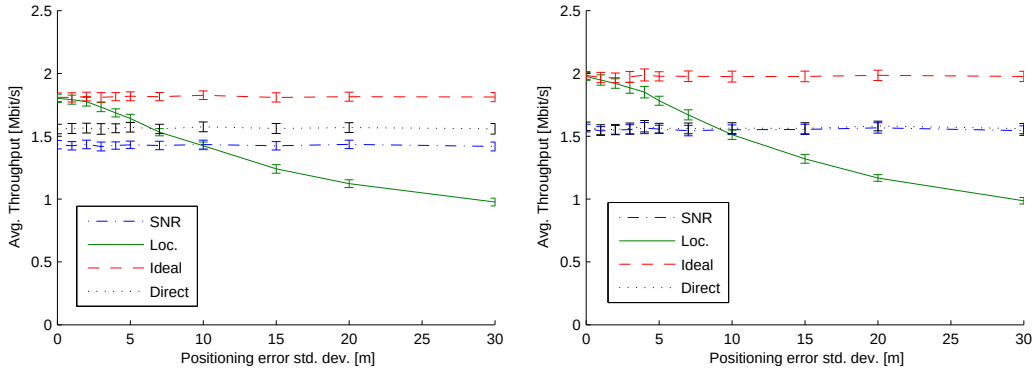


Figure 5.15: BER impact of the accuracy of location measurements.

crossing $\sigma_{\text{pos}} \simeq 13m$. Considering that GPS receivers typically achieve an accuracy of $15m$ [Garmin, 2009], a GPS-only localization system may not be accurate enough for location based relaying. A solution would be to consider a hybrid localization approach using both GPS, Galileo and network based localization techniques that have been shown to improve location accuracy as described in chapter 2 and chapter 4.



(a) BER based relay selection (b) Throughput based relay selection

Figure 5.16: Throughput impact of varying location error.

Figure 5.16 shows the corresponding results where the performance metric is throughput, and the relay selection criteria is BER or throughput based. Both results are quite similar to the BER results in Figure 5.15, in the sense that the location based scheme is close to the ideal scheme for $\sigma_{\text{pos}} < 5m$. However in Figure 5.16(a) the location based scheme crosses the direct scheme at $\sigma_{\text{pos}} \simeq 7m$, and in Figure 5.16(b) it crosses at $\sigma_{\text{pos}} \simeq 9m$.

Varying path loss exponent guess

Since the path loss exponent cannot be assumed to be known, we investigate the impact of varying the guessed value of n in Figure 5.17. Interestingly, the performance of the location based scheme is very close to the ideal scheme for values in a relatively wide range of $2 < n < 4$. Correct estimation of the path loss exponent does not seem to be highly important for achieving a near ideal performance with the location based scheme.

Considering the corresponding throughput plots in Figure 5.18, we see that the BER based selection results in a nearly similar performance for wide range of path loss exponent values. The throughput based selection yields higher performance than the BER based for values close to the true value (approximately

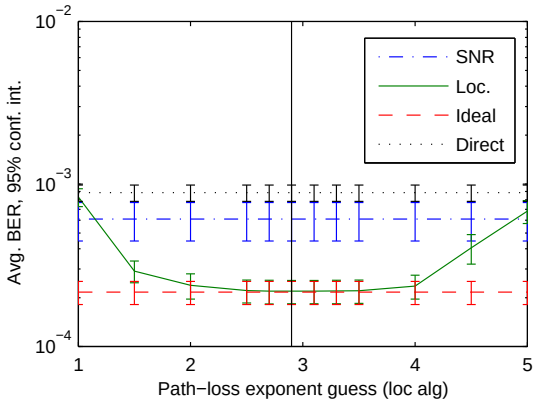
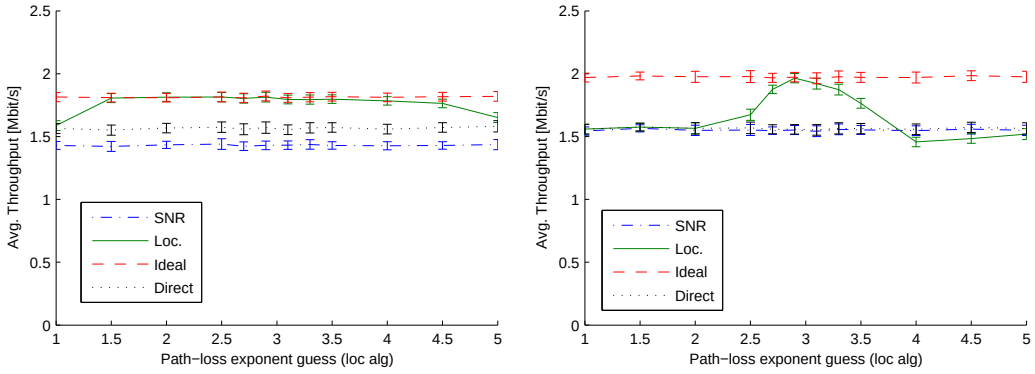


Figure 5.17: BER impact of the accuracy of the used path loss exponent compared to the true value of 2.9.

±0.3), but performs worse for values outside this range. The throughput based selection is clearly more sensitive to inaccurate estimates of the path loss exponent than the BER based selection.



(a) BER based relay selection (b) Throughput based relay selection

Figure 5.18: Throughput impact of the accuracy of the used path loss exponent compared to the true value of 2.9. Node speed is 5 m/s.

Varying NLOS attenuation

There may be cases where the direct propagation path between two nodes is blocked by obstacles. This NLOS condition may occur between two MDs or the AP and an MD. In this work we have introduced a horizontal "wall" that attenuates all crossing transmissions, but does not hinder node movements. In Figure 5.19 we investigate the impact of varying the wall attenuation for a

wall that is placed $0.25m$ below the AP. That is, the AP has LOS to all MDs in the upper half of the scenario and NLOS towards all MDs in approximately the lower half.

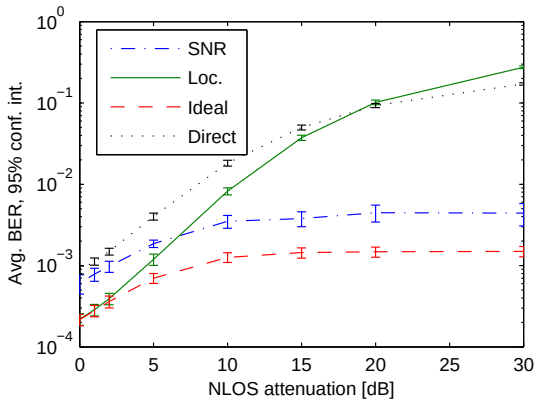
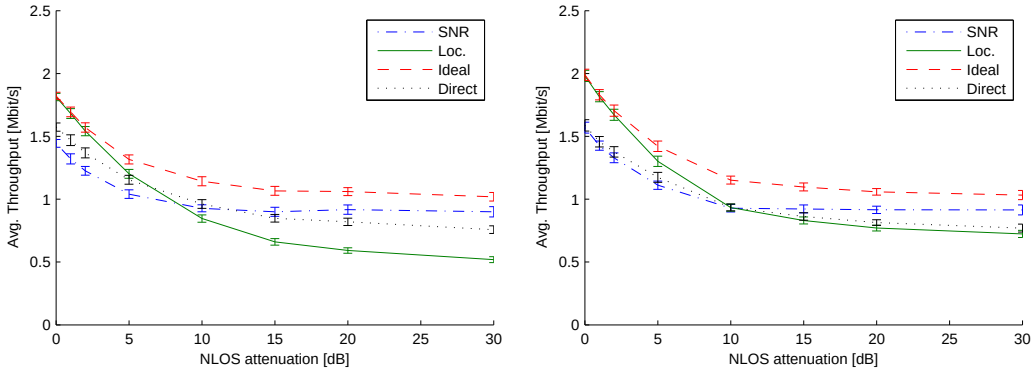


Figure 5.19: BER impact from NLOS caused by horizontal wall for increasing attenuation.

All schemes achieve worse performance for increasing wall attenuation. However, it is clear that the ability to sense the wall attenuation is beneficial. This is clear by noticing that the performance of the ideal and SNR based schemes does not degrade as much as the location based and direct schemes. When transmitting to an MD in the lower half of the scenario, the direct scheme will experience the attenuation for all transmissions from the AP to such nodes, whereas the ideal and SNR based schemes can take the attenuation into account when selecting a relay node. Since the location based schemes uses only the path loss model for predicting link states, the wall attenuation is not taken into account and the relay can even be chosen in such unfortunate way that the wall is crossed in both the AP-relay and relay-destination transmissions. This can be seen in Figure 5.19, where the BER of the location based scheme even exceeds the direct scheme.

The plots in Figure 5.20 shows how the BER and throughput based schemes cope with unknown NLOS attenuation. The most notable difference is the impact on the location based scheme. For this scheme, the throughput based selection does not suffer so much from the unknown attenuation as the BER based selection does. Generally, the plots show that even a little unknown attenuation has a big negative impact on throughput performance.



(a) BER based relay selection (b) Throughput based relay selection

Figure 5.20: Throughput impact of increasing NLOS attenuation.

Varying NLOS attenuation guess

Assuming that the AP has access to a spatial map of obstructions that cause NLOS conditions, the performance of the location based scheme can be improved in NLOS conditions. The map could be used to determine if there is LOS between two node positions. For LOS situations the attenuation term X in (5.6) is zero, while for NLOS it will be nonzero. But as the wall attenuation may very well be unknown, we investigate the performance impact of different guesses for the attenuation value. We use a wall attenuation of $13.3dB$ in our simulations, since according to reference [Durgin et al., 1998] this value is typical for home walls.

Figure 5.21 shows that as the NLOS attenuation guess gets close to the true value of $13.3dB$, the avg. BER of the location based scheme decreases. Guesses exceeding the true value do however not cause an increase in the BER.

The throughput plots in Figure 5.22 basically show the same behavior, and there is no noteworthy difference between the BER and throughput selection schemes. Notice that the transmit power has been increased to 300 mW, as the NLOS attenuation resulted in very low throughput.

In addition to the wall position $0.25m$ below the AP, we also investigate the situation where the wall is half-way between the AP and the bottom border of the scenario, which corresponds to $25m$ below the AP. This result is shown in Figure 5.23.

Primarily, we can see that the conditions have become difficult for even the ideal scheme as the BER has shifted a decade up, compared to Figure 5.21.

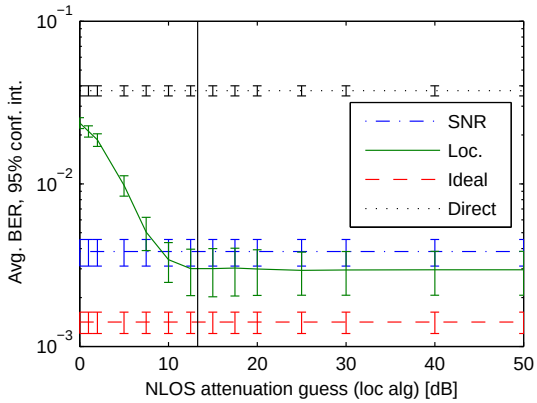
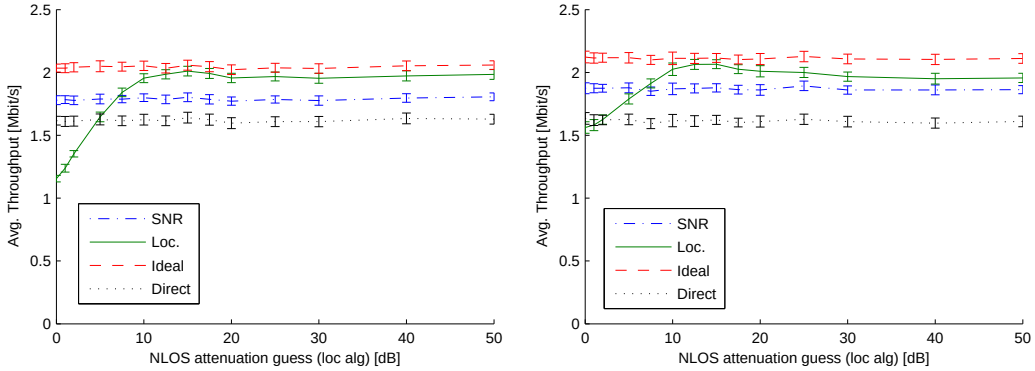


Figure 5.21: BER impact for different guesses of the unknown NLOS attenuation with a true value of $13.3dB$. Wall is close to AP. Node speed is $5m/s$.



(a) BER based relay selection (b) Throughput based relay selection

Figure 5.22: Throughput impact for different guesses of the unknown NLOS attenuation with a true value of $13.3dB$. Wall is close to AP. Tx power is 300 mW.

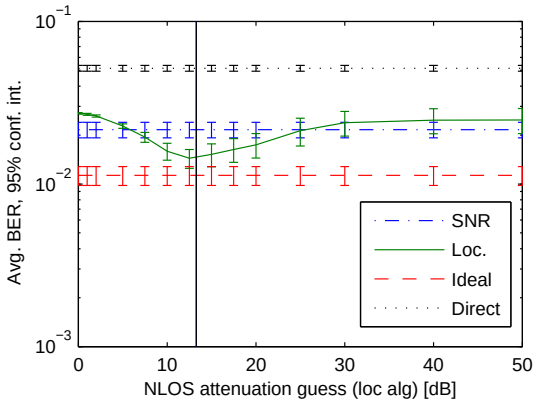
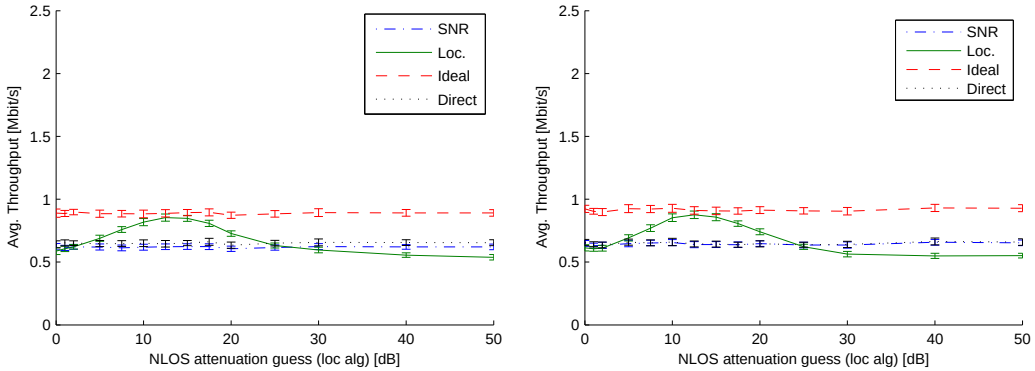


Figure 5.23: BER impact for different guesses of the unknown NLOS attenuation with a true value of 13.3dB. Wall is halfway between AP and bottom.

Further, it can be seen that under- or overestimation of the wall attenuation has a clearly negative impact on performance. So in this case, a priori knowledge of the wall attenuation is needed to obtain good performance. The wall attenuation could be obtained by evaluating both SNR and location measurements in the online system, however this is a topic for future work.



(a) BER based relay selection (b) Throughput based relay selection

Figure 5.24: Throughput impact for different guesses of the unknown NLOS attenuation with a true value of 13.3dB. Wall is halfway between AP and bottom. Tx power is 300 mW.

Finally, the plots in Figure 5.24 shows the corresponding throughput plots. Like the BER had increased for this wall position compared to the previous wall position, the throughput has gone down compared to Figure 5.22. The BER and throughput based selection schemes are equally affected by inaccurate guesses of the NLOS attenuation. Within a range of $\pm 5dB$, the achieved

throughput is close to the ideal scheme, which means that $\pm 5dB$ is an acceptable deviation.

5.6.4 Outlook on Analytic Model

In the results presented above, the location based scheme has been using a high update frequency for location information. While this was a fair assumption for the comparison to the SNR based scheme, as they used the same signaling overhead, in practice it may be useful to have less frequent location updates and spend only the necessary signaling overhead. The presented evaluation results and analysis framework are clearly useful for choosing meaningful values of the system parameters related to measurement update frequency and storage threshold. However, it can be a cumbersome task to simulate all imaginable parameter combinations in order to set the system parameters correctly. A very useful extension of this work would therefore be to consider an analytic model, which could be used to determine the system parameters more efficiently.

A starting point for constructing such an analytic model could be to focus on the decision of whether to use a given relay node for a source and destination node pair or not, as sketched in Figure 5.25. The idea is based on the fact that for a source and destination node pair, the usefulness of a relay depends on the position of the relay. The closer the relay node is to the center of the gradient circle in the figure, the better suited it is. Of course this also depends on the positions of the source and destination nodes. If they can already use the highest possible bit rate for direct transmissions, there is no need to use the relay. Notice that the benefit-gradient may not necessarily be a circle as sketched in the figure.

In order to decide if a relay is useful, the envisioned analytic model should be able to give an estimate of the achievable performance gain of using the relay node. The idea is to have the model take into account the age of measurements, the node movement speeds, and the localization error, when calculating the estimated gain. These factors are shown in the figure as uncertainty regions around the mobile relay node and the mobile destination node. The meaning is that actual positions of these two nodes are outlined by these uncertainty regions. Assuming that the AP, which uses the model, has a set of measurements of the relay and destination nodes' positions. Now, given the age of these measurement and the nodes' movement speed, the probability distribution of their current positions could be computed. Further, since the location measurements are most likely inaccurate to some degree, the probability distribution of the location error should also be taken into account. In the figure,

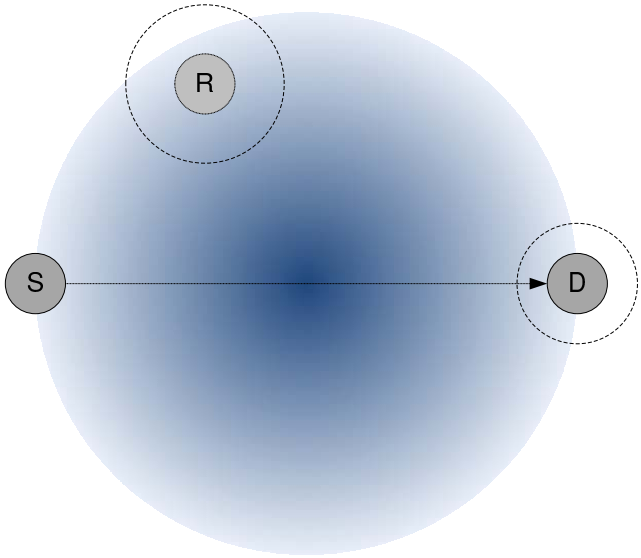


Figure 5.25: Principle of analytic model for relay selection. The gradient shows the benefit of different relay positions (darker is better), and the dashed circles are uncertainty regions, due to localization error, measurement age and node mobility.

these probabilistic uncertainties have been illustrated in a simplified way with the dashed circles. In the model, the probability distributions should be used. If the movement directions of the nodes are unknown, an assumption of all angles being likely possible would probably be useful. On the other hand, if movement prediction was used, an approximated movement direction could reduce the uncertainty of the model. A possible application of the model would then be to determine if a given relay would be beneficial to use with, e.g., 95 % confidence.

A possibly framework for this analytic model would be the Mismatch Probability (mmPr) framework described in [Bogsted et al., 2010]. The uncertainty regions described above, can be formulated as the mismatch probability of location information after a probabilistic delay, which can be calculated analytically using this framework, as it is done in [Olsen et al., 2010]. This work would then need to be extended to take into account the relaying aspect.

5.7 Conclusion

The scenario considered in this work concerns downlink data transmissions in a IEEE 802.11 based wireless network with mobile users, where transmissions can be either direct or two-hop relayed. The focus of this work has been on analyzing the impact of mobility, delays and inaccurate information on an SNR measurement based and a location based relay selection scheme. The schemes have been evaluated using ns-2 and matlab simulations for the random waypoint mobility model. The results of the two relaying schemes were compared to a scheme that always uses direct transmissions and an ideal scheme that has instant and perfect link state knowledge. Also, both a BER based and a throughput based relay selection criteria were considered.

The results show that for relatively fast moving mobile devices ($5-15\ m/s$) the achieved avg. BER performance of the SNR measurement based scheme can get significantly worse than the always direct scheme. Increasing the hello broadcast rate can mitigate this effect, however, at the cost of a linear increase in signaling overhead. We show that by limiting the measurement storage time, i.e. letting the AP use only fresh measurements, we are able to achieve a performance slightly better than or equal to the always direct scheme without increasing the signaling overhead. This result underlines the importance of choosing the storage threshold parameter carefully in scenarios with mobile devices.

In cases with fast moving users, the measurement collection frequency that is required for acceptable performance results in a large signaling overhead. In these cases the location based scheme, which uses a path loss model for relay selection, creates considerably less signaling overhead than the SNR measurement based scheme. We have shown that due to reduced signaling overhead, this scheme allows considerably higher movement speeds as compared to the SNR measurement based scheme. In the considered case, a four fold increase of the movement speed was possible.

Further we found that the location accuracy required to achieve close to ideal performance was around 3-4 m standard deviation. With the considered localization system in Chapter 4 having an approximated standard deviation of 2-3 m, it can be concluded that location based relaying is feasible with this kind of system. However, for the GNSS and cellular based localization systems presented in Figure 2.5 on page 27 the estimated standard deviations of 8-13 meters would not be acceptable.

As the parameters of the environment are not always known in advance, we have investigated the sensitivity of the location based relaying scheme towards

inaccurate settings of parameter in the path loss model. With regard to the path loss exponent, which is typically either unknown and thus guessed from the environment characteristics or estimated as the average over a larger area, we found that estimates within a relatively wide range of ± 1.5 around the true value resulted in near-optimal results for the BER relay selection criteria. This was significantly different for the throughput critera selection, as the acceptable range here was only in the order of ± 0.3. In the case of a NLOS situation, we found that the relaying performance was severely degraded if the location based relaying did not have a priori knowledge of the NLOS situation. If on the other hand knowledge of LOS/NLOS between nodes was made available by extending the location based scheme with a spatial map of obstructions, the obtained performance was useful for estimates within $\pm\ 3dB$ of the true attenuation factor. A hybrid scheme that combines the low overhead from the location based scheme with the sensing ability from the SNR based scheme would therefore be an interesting topic for future work. Specifically, location information could be used for selecting a few of the most promising relay candidates, for which dedicated link measurements are obtained in order to detect and adapt to the actual channel state on those links.

Generally, it was noticed that the throughput selection criteria yields slightly higher throughput than the BER selection criteria, whereas the BER criteria seems to be more robust, for example towards inaccurate path loss exponent values.

A promising future work item would be to develop an analytic model which is able to judge the usefulness of different relay choices when taking into account the possible movements, localization inaccuracy, and measurement age. This could be realized by extending the work in [Olsen et al., 2010] to include the relaying aspects.

References

M. Bogsted, R.L. Olsen, and H.P. Schwefel. Probabilistic models for access strategies to dynamic information elements. *Performance Evaluation*, 67(1): 43–60, 2010. ISSN 0166-5316.

Q. Chen, F. Schmidt-Eisenlohr, D. Jiang, M. Torrent-Moreno, L. Delgrossi, and H. Hartenstein. Overhaul of IEEE 802.11 modeling and simulation in ns-2. *Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems*, 2007.

G. Durgin, TS Rappaport, and H. Xu. Measurements and models for radio path loss and penetration loss inand around homes and trees at 5.85 GHz. *IEEE Transactions on Communications*, 46(11):1484–1496, 1998.

Garmin. About gps, 2009. URL `http://www8.garmin.com/aboutGPS/`.

Zhengqing Hu and Chen-Khong Tham. Ccmac: Coordinated cooperative mac for wireless lans. *Computer Networks*, 54(4):618 – 630, 2010. ISSN 1389-1286. doi: 10.1016/j.comnet.2010.02.001. Advances in Wireless and Mobile Networks.

IEEE. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999)*, pages C1–1184, 12 2007.

J.-S. Liu and Y.-C. Wong. A relay-based multirate protocol in infrastructure wireless lans. In *Networks, 2004. (ICON 2004). Proceedings. 12th IEEE International Conference on*, volume 1, pages 201 – 206 vol.1, 2004. doi: 10.1109/ICON.2004.1409124.

P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S.S. Panwar. CoopMAC: A cooperative MAC for wireless LANs. *IEEE Journal on Selected Areas in Communications*, 25(2):340, 2007.

M.H. Lu, P. Steenkiste, and T. Chen. Design, implementation and evaluation of an efficient opportunistic retransmission protocol. In *Proceedings of the 15th annual international conference on Mobile computing and networking*, pages 73–84. ACM, 2009.

S. Narayanan and S.S. Panwar. To Forward or not to Forward–that is the Question. *Wireless Personal Communications*, 43(1):65–87, 2007.

R.L. Olsen, J. Figueiras, J.G. Rasmussen, and H.P. Schwefel. How precise should localization be? A quantitative analysis of the impact of the delay and mobility on reliability of location information. *Proc. of IEEE GLOBECOM*, 2010.

J.G. Proakis. Digital Communications Third Edition. *McGrawHill Inc, New York, USA*, 1995.

K. Tan, Z. Wan, H. Zhu, and J. Andrian. CODE: cooperative medium access for multirate wireless ad hoc network. In *IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, 2007.

Bin Zhao and M.C. Valenti. Practical relay networks: a generalization of hybrid-arq. *IEEE Journal on Selected Areas in Communications*, 23(1):7–18, 2005.

H. Zhu and G. Cao. rDCF: A relay-enabled medium access control protocol for wireless ad hoc networks. *IEEE Transactions on Mobile Computing*, pages 1201–1214, 2006. ISSN 1536-1233.

M. Zorzi and R.R. Rao. Geographic random forwarding (GeRaF) for ad hoc and sensor networks: multihop performance. *Mobile Computing, IEEE Transactions on*, 2(4):337–348, 2004. ISSN 1536-1233.

6

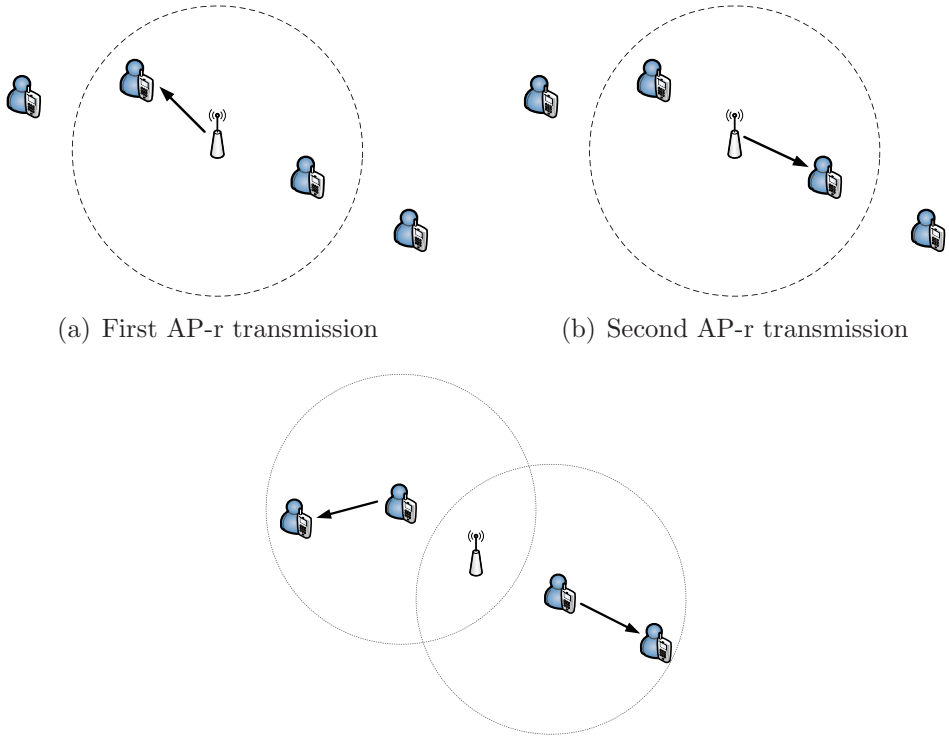
Joint Location based Relay Selection and Power Adaptation

The contribution of this chapter is an extension of the centralized relay selection mechanisms presented in the previous chapter. Specifically, a cross-layer optimization is proposed, which uses location information to jointly choose the optimal relay and relay transmit powers that allow two concurrent relay-to-destination transmissions. The outcome of this optimization is an increased throughput for certain node constellations compared to regular two-hop relaying. As this optimization relies on location information for interference prediction, the impact of inaccurate location information is analyzed.

6.1 Introduction

In the previous chapter 5, the two-hop relaying technique is used to improve throughput for nodes located near the edge of the AP coverage zone in the small-scale scenario. Instead of always using a direct transmission from the AP to the destination node, an intermediate relay node is used, since this can improve performance, as the relay node is typically closer to the destination than the AP.

Just like spatially well-separated Wi-Fi hotspots can work simultaneously it may therefore be possible to have spatially separated relay-to-destination (r-d) transmissions ongoing simultaneously, as sketched in Figure 6.1. In such a case it could be possible to increase the total throughput, if the transmission time can be lowered. A prerequisite for this is however that the mutual interference that each r-d transmission creates does not disturb the other r-d transmission so much that the frame loss leads to a lower throughput.



(a) First AP-r transmission

(b) Second AP-r transmission

(c) Simultaneous r-d transmissions

Figure 6.1: Principle of simultaneous relay to destination transmissions.

The research question being addressed in this chapter is therefore to investigate if location information can be used to allow simultaneous r-d transmissions, especially how suitable location information is for interference prediction, and how accurate the location information needs to be. Since the level of mutual interference depends on the positions of the relay nodes, destination nodes, and the used transmit power, these parameters should be taken into account. More specifically, this chapter concerns a location based cross-layer optimization that determines the best relay nodes and relay transmit powers for a pair of destination nodes, in terms of maximizing throughput when taking into account the generated mutual interference. In the following, this scheme will be referred to as SimTX, which is an acronym for simultaneous transmissions.

Figure 6.2 relates this contribution to the contribution in the previous chapter 5. Here, interference is taken into account for selecting relays and relays' transmit powers. The greyed out parts in the figure, namely measurement collection and error terms, are not considered in detail in this chapter but are accounted for in combination as a general localization error.

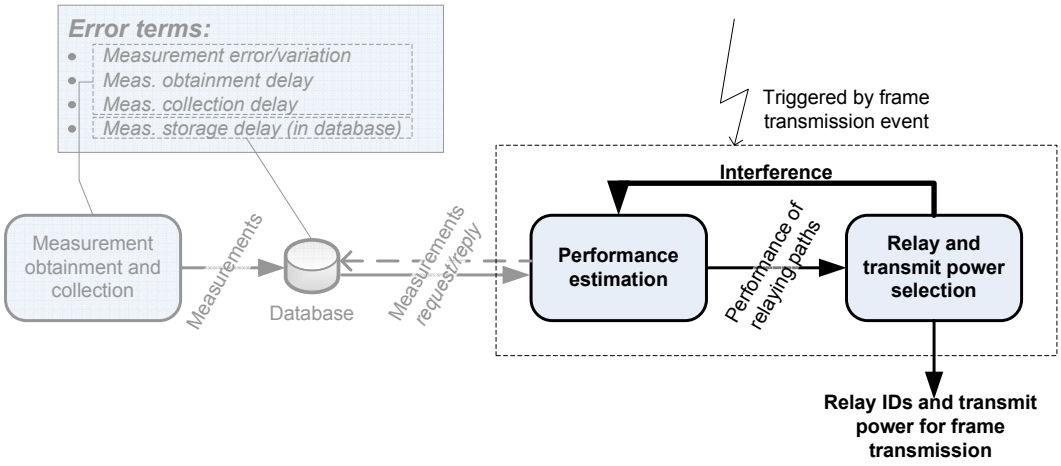


Figure 6.2: Extended version of the general relaying system in Figure 5.2. The grayed-out parts are not considered in details in this chapter, whereas the bold text shows the additions.

Since the principle of simultaneous transmissions (sketched in Figure 6.1) requires two destination nodes, it is necessary to compare the achieved results to the case of non-simultaneous transmissions to the same two destination nodes. For evaluation, we compare the achieved performance of the SimTX scheme to existing transmissions schemes, namely direct transmission and sequential two-hop relayed transmission, as described in chapter 5.

An example of the possible node layout for the considered small-scale scenario is shown in Figure 6.3. Here, the two destination nodes, denoted primary and secondary, are shown in solid, and the corresponding candidate relays are shown with plusses and crosses. A detailed description of the applied relay selection scheme is given later in section 6.3. For simplicity we only consider downlink transmissions initiated from the AP to mobile nodes, and do therefore not consider potential influence from mobile nodes in surrounding networks.

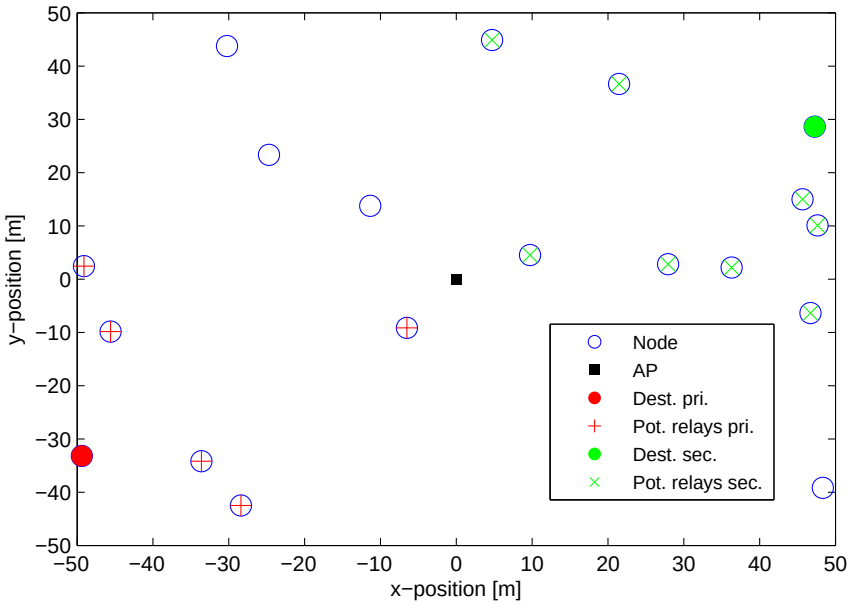


Figure 6.3: Example of the node layout, with primary and secondary destination nodes and their potential relays.

6.2 Related Work

The idea of interference aware tuning of transmit power, to allow co-existing users, is well known in the literature. This section presents some relevant examples of related work.

General DCF-optimizations

In much existing work related to IEEE 802.11, the RTS/CTS scheme is used to obtain CSI, which is then exploited for transmit power and rate adaptation. One such example is [Qiao et al., 2003] that presents an energy-effective DCF.

Here the RTS/CTS handshake is used to probe the link quality and from this select the most appropriate PHY mode and transmit power level in terms of delivered data per joule.

The DCF extension proposed in [Chevillat et al., 2005] does transmit power and rate adaptation, only from the successful or unsuccessful reception of an ACK at the source. If an ACK is not received, the source concludes that the signal quality was insufficient and it increases the transmit power or lowers the data rate. If several ACKs are received successfully for several transmissions, the transmit power is lowered or the data rate is increased.

In [Nadeem et al., 2005] a similar extension of the 802.11 DCF is proposed. However, instead of probing the link quality, location information is included in the RTS/CTS handshake. By using a propagation model, the nodes are able to make better interference predictions and blocking assessments, which increases the achieved throughput.

Ad hoc networks

Another approach, targeted at ad hoc networks is presented in [Kim et al., 2006]. The authors achieve capacity and performance improvements of up to 22 % by interference-aware tuning of transmit power, bit rate, and carrier sense threshold.

Further, in [ElBatt and Ephremides, 2004] the authors propose a scheme for joint scheduling and power control for wireless ad hoc networks. The algorithm works in two steps, where in the first step the scheduling algorithm coordinates transmissions of independent users so that strong levels of interference are avoided, and in the second step the power is allocated to transmissions.

Infrastructure oriented

As the AP in an infrastructure oriented network is the central data exchange point, it is often a bottleneck and thus an obvious case for optimization. The following examples of existing work are targeted at such scenarios.

Initially, in [Lim et al., 2006], the authors improve downlink throughput near Internet gateways by using signal strength measurements to create a virtual interference map, from which different multihop routing paths are planned, thereby allowing simultaneous transmissions.

Another scheme, which improves throughput in 802.11 WLANs by means of simultaneous transmissions, is the CCMAC protocol proposed in [Hu and Tham, 2010]. The protocol uses a partially observable Markov decision process model to determine the best transmission rate and relay nodes. Further, it allows simultaneous source-to-relay transmissions for uplink transmissions to the AP. For selecting relay nodes and determine usable bit rates, the scheme uses past link measurements and overheard transmissions, similarly to the relaying protocol CoopMAC [Liu et al., 2007].

In [Ma et al., 2008] the authors propose a transmit power and carrier sense threshold adaptation algorithm for high density IEEE 802.11 WLANs (both infrastructure and ad hoc). The algorithm detects the type of frame loss as either interference or collision, and based on that it jointly tunes the corresponding network parameters (physical carrier sensing threshold, transmit power) accordingly. They show that in high density WLANs the link throughput is greatly increased compared to using only carrier sensing threshold adaptation.

Finally, in [Li et al., 2010] the authors propose a differential evolution based algorithm for power control and scheduling in wireless mesh networks. By foreseeing possible conflicts based on the traffic patterns across the network, the algorithm is able to calculate the optimal active time of each traffic pattern as well as bit rate and power level for the corresponding relay nodes.

In the reviewed literature there are examples of location based power adaptation, joint relay selection and power adaptation, and interference aware power adaptation for allowing simultaneous transmissions. However, to the best of our knowledge there are no previous examples in the literature of "centralized location based joint relay selection and power adaptation for simultaneous downlink transmissions".

6.3 The SimTX Scheme

The sequence of events in the case where two r-d transmissions occur simultaneously is sketched in Figure 6.4. Initially, the AP runs the cross-layer optimization algorithm, which determines the relays and relay transmit powers to be used for the primary and secondary transmissions. This calculation is based on information regarding the positions of the AP and the mobile nodes, which is collected periodically as described in the previous chapter 5. Hereafter, the two AP-r transmissions are done sequentially using the maximum transmit power level. For the simultaneous r-d transmissions we assume that the relay nodes can choose a power level from a fixed and discrete set. The aim of the

adjustments is to limit cross-interference. We assume that the AP is able to schedule transmissions in a way so that R1 and R2 can transmit simultaneously to D1 and D2, respectively, and that the MAC protocol is modified so that this is possible. These simultaneous transmissions happen immediately after the ACK for Data2 has been transmitted by R2 to the AP. If R1 does not overhear the ACK from R2, it should wait the time of an ACK, a DIFS, and a short back-off period, to allow a possible retransmission of Data2. When R1 overhears the ACK from R2, it means that R2 has successfully received Data2, and both R1 and R2 are ready to start the simultaneous transmissions. R1 and R2 should therefore start their transmissions immediately hereafter. This can be ensured if R1 and R2 initiate the simultaneous transmissions after waiting for the time of a SIFS, after the ACK from R2 has been transmitted. This behavior would require a modification to the MAC protocol, so that relays can be instructed to act in two additional transmission modes: two-hop relayed or simultaneous transmissions. In practice, the AP would use additional frame header fields in the data frames to notify the relays of the transmission mode and transmit power to use for a given pair of transmissions.

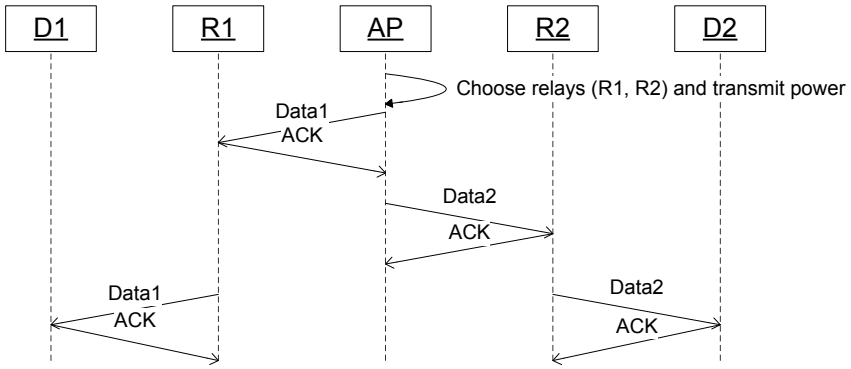


Figure 6.4: D1 and D2 are the destination nodes, R1 and R2 are the relays and AP is the source of the transmissions.

The proposed cross-layer algorithm for joint relay selection and power adaptation works as a scheduling algorithm for the AP's transmit queue. In order for two packets to be suited for simultaneous transmission, their destination nodes need to be located so that the simultaneous r-d transmissions do not interfere too severely. Assuming that the AP knows the positions of all relay and destination nodes, it can evaluate the expected interference for specific choices of destination pairs. The task of the relaying scheme is therefore first to schedule the packets in the queue in such a way that pairs of destination nodes that are suited for simultaneous transmissions are scheduled accordingly,

and further to select the most suitable relay nodes and transmit power levels. For this we propose an algorithm which is divided into the following steps:

Form pairs The AP picks the first packet in the transmit queue as the primary destination node. In order to have simultaneous transmissions a secondary destination node is also needed. We assume that a scheduled transmission can be upgraded to be the secondary transmission if it seems suited for simultaneous transmission. This is similar to the approach used in [Lim et al., 2006]. To limit cross-interference, we select the secondary node as the node that is closest to the coordinate: $(-x_{\text{pri}}, -y_{\text{pri}})$.

Since the choice of secondary destination node is not independent of the relay positions, another node could be more suited. However, to keep complexity low, we do not consider the destination node selection as a part of the optimization problem.

Choose potential relays Having determined the two destination nodes, the best suited relay nodes are sought. Initially, a pre-filtering is performed to rule out unsuited relays. As potential relays we consider those where both the *AP-r* distance and the *r-d* distance are less than the *AP-d* distance:

$$d_{\text{AP-d}} > d_{\text{AP-r}} \quad \wedge \quad d_{\text{AP-d}} > d_{\text{r-d}}. \tag{6.1}$$

Find max-throughput configuration In this step, the algorithm solves the 4-dimensional optimization problem of finding the best configuration of relays and transmit power. It does this by computing the expected throughput of all combinations of primary and secondary candidate relays and available transmit power levels for these relays. The best configuration is the combination that has the highest throughput.

Transmit data frames Given the determined best configuration, the AP now transmits the data frames either directly to the destination nodes or to the relay nodes. The transmit power level to be used by the relay nodes, should be specified by the AP as an extra option in the frame header.

6.4 Throughput Calculation

In order for the SimTX algorithm to evaluate the expected throughput of different relay and transmit power configurations, we apply the throughput model presented in chapter 3. As we are interested in comparing the performance of the proposed SimTX relaying scheme to existing transmission schemes, we consider also the expected throughput for direct transmissions as well as two-hop relaying.

For comparison we calculate the average throughput, which is $\frac{\text{Delivered data}}{\text{Transmission time}}$, of the primary (pri) and secondary (sec) transmissions of the considered algorithms. For the direct algorithm the throughput[1] is:

$$S_{\text{dir}} = \frac{(P_{\text{suc}}^{\text{pri}} + P_{\text{suc}}^{\text{sec}}) \cdot B_{\text{MSDU}}}{E[T_{\text{tx}}^{\text{pri}}] + E[T_{\text{tx}}^{\text{sec}}]} \tag{6.2}$$

where the MAC payload size B_{MSDU} is given in octets.

In the following, we use the indices 1 and 2 to indicate the AP-r and r-d transmissions. The throughput for the two-hop relaying algorithm is calculated as:

$$S_{\text{rel}} = \frac{(P_{\text{suc}}^{\text{pri},1} P_{\text{suc}}^{\text{pri},2} + P_{\text{suc}}^{\text{sec},1} P_{\text{suc}}^{\text{sec},2}) \cdot B_{\text{MSDU}}}{E[T_{\text{tx}}^{\text{pri},1}] + E[T_{\text{tx}}^{\text{pri},2}] + E[T_{\text{tx}}^{\text{sec},1}] + E[T_{\text{tx}}^{\text{sec},2}]}. \tag{6.3}$$

Finally, we calculate throughput for the SimTX algorithm. In this equation, the parameters with index 2, i.e. corresponding to the simultaneous r-d transmissions, are based on the SINR calculation in equation (3.2) where the mutual interference is taken into account. The resulting throughput is:

$$S_{\text{sim}} = \frac{(P_{\text{suc}}^{\text{pri},1} P_{\text{suc}}^{\text{pri},2} + P_{\text{suc}}^{\text{sec},1} P_{\text{suc}}^{\text{sec},2}) \cdot B_{\text{MSDU}}}{E[T_{\text{tx}}^{\text{pri},1}] + E[T_{\text{tx}}^{\text{sec},1}] + E[\max(T_{\text{tx}}^{\text{pri},2}, T_{\text{tx}}^{\text{sec},2})]}. \tag{6.4}$$

$E[\max(T_{\text{tx}}^{\text{pri},2}, T_{\text{tx}}^{\text{sec},2})]$ is calculated from the cdf of the maximum of two independent RVs X and Y:

$$\begin{aligned} P(\max(X,Y) \leq c) &= P(X \leq c \text{ and } Y \leq c) \\ &= P(X \leq c)P(Y \leq c) = F_Y(c)F_X(c). \end{aligned} \tag{6.5}$$

This is true for independent RVs. However, if one transmission is successful and the other fails, the single retransmission experiences a better SINR. The

[1] To get throughput in Mbit/s, multiply the results by $8 \cdot 10^{-6}$.

assumption of independence between $T_{\text{tx}}^{\text{pri},2}$ and $T_{\text{tx}}^{\text{sec},2}$ is therefore expected to result in slightly pessimistic results. The influence of this assumption is investigated by simulation in section 6.6. Consequently we get:

$$F_{T_{\text{tx}}^{(2)}}(t) = P(\max(T_{\text{tx}}^{\text{pri},2}, T_{\text{tx}}^{\text{sec},2}) \leq t) = F_{T_{\text{tx}}^{\text{pri},2}}(t) F_{T_{\text{tx}}^{\text{sec},2}}(t) \tag{6.6}$$

where $F_{T_{\text{tx}}^{(2)}}(t)$ is the product of the cdfs of the time spent per transmission attempt on the simultaneous r-d transmissions, which is simply the elementwise product of two vectors of length $R+1$. Since $T_{\text{tx}}^{(2)} > 0$, we may compute the expectation of the maximum as:

$$E[\max(T_{\text{tx}}^{\text{pri},2}, T_{\text{tx}}^{\text{sec},2})] = \int_0^\infty (1 - F_{T_{\text{tx}}^{(2)}}(t)) dt. \tag{6.7}$$

6.5 Evaluation Methodology

The proposed SimTX scheme is evaluated through extensive simulations. The simulation framework used for this has been programmed in matlab. The Wi-Fi protocol considered in this work, is assumed to be based on IEEE 802.11a PHY and MAC layers described in [IEEE, 2007], however modified so that the AP can do frame re-scheduling, and instruct relay nodes on which transmission mode to use and which transmit power to use, through custom frame header fields. For simplicity it is further assumed that all transmission are of the same payload length and that all entities use the same fixed modulation and coding scheme, which is the 6 Mbit/s mode specified for IEEE 802.11a. Also, the nodes are assumed to be statically positioned (i.e., not moving around) and their locations are assumed to be known perfectly.

Also in this scenario it is assumed that there is no inter-cell interference from neighboring cells. Since IEEE 802.11a supports 11 independent channels, this assumption is considered reasonable in situations where channels have been allocated properly.

Specifically, for each simulation run, the following is performed:

1. Generate random node layout
2. Randomly select primary destination node
3. Run the algorithm to find secondary destination node and candidate relays, and for each of the considered transmission modes (direct, two-hop, and SimTX), calculate the best relay and transmit power configuration, in terms of expected throughput.

4. Calculate the achieved throughput for the best configuration for each mode, given the known node locations. These are the evaluation results.

In both the third and fourth step, the throughput is calculated using the formulas in section 6.4. In the third step, the three transmission modes are treated differently. For the direct scheme, only the two (sequential) direct transmissions are possible. For the two-hop relaying mode, the expected throughput for all relay candidates for each destination node is calculated, and for each destination node, the relay, which leads to the highest expected throughput is selected. These two relays constitute the best configuration. The number of combinations that are considered is simply: $N_{\text{cand}}^{\text{D1}} + N_{\text{cand}}^{\text{D2}}$, where $N_{\text{cand}}^{\text{D}}$ is the number of candidate relays for destination node D. For the SimTX algorithm, the candidate relays for each destination node cannot be considered independently as for the two-hop relaying mode, since they are influencing each other during the simultaneous transmission phase. Therefore, it is necessary to iterate over all combinations of relays among the two sets of relay candidates and all combinations of relay transmit power levels. This amounts to the following number of combinations:

$$N_{\text{cand}}^{\text{D1}} \cdot N_{\text{cand}}^{\text{D2}} \cdot (N_{\text{Ptx}})^2 \tag{6.8}$$

where $N_{\text{cand}}^{\text{D}}$ is the number of candidate relays for destination node D and N_{Ptx} is the number of available transmit power levels for each relay node.

In order to validate the independence assumption used to calculate the expected throughput for the SimTX algorithm in the previous section 6.4 and the simplifying assumption of the ACK having the same SINR as the foregoing data transmission, an embedded simulation of the IEEE 802.11 back-off mechanism has been created. This embedded simulation simulates the transmission and reception of the individual data and ACK frames for the AP-to-relay and the simultaneous relay-to-destination transmissions. The outcome of each frame transmission is stochastically determined based on the frame size and the BER. Here the BER level depends on the SINR level for the particular transmission. If for example one of the destination nodes in a simultaneous relay-to-destination transmission does not successfully receive the data frame but the other destination does, the retransmission of the failed transmission will have a higher SINR, since there is no interference. These embedded stochastic simulations were repeated 10 000 times.

For all results shown in the following section 6.6 but the last, the locations of the relay and destination nodes are assumed to be perfectly known. However, for the results in Figure 6.11, a stochastic zero-mean two-dimensional Gaussian distributed location error has been added to each of the relay and

destination nodes' locations, in order to investigate the impact of inaccurate location information. In this case, the SimTX algorithm is using the inaccurate location estimates for determining the best transmission mode (direct, two-hop, or SimTX) and configuration (relays and transmit power). In order to evaluate the achieved throughput, the chosen configuration is therefore applied to the equations in section 6.4 in combination with the actual locations of the relay and destination nodes.

6.6 Results and Discussion

In this section we use the simulation framework described in the previous section to evaluate the proposed SimTX algorithm in comparison to direct transmissions and two-hop relaying. The simulation study considers a small-scale scenario, specified in details by the parameters in Table 6.1.

Parameter	**Value**
Scenario size	100m x 100m
AP position	(0,0)
Number of nodes	30
Path loss exponent (α)	2.9
$N0_{\text{floor}}$	$-86dBm$
Ricean K-factor (K)	15
Bit rate	6 Mbit/s
Modulation scheme	BPSK
Max no. of retransmissions (R)	7
B_{MSDU}	1024 bytes
P_{tx} levels available	0, 5, 10, 20, ... , 90, 100 mW

Table 6.1: Scenario parameters

The following results are calculated using the throughput model in the scenario specified in Table 6.1. The BER-throughput relationship of the throughput model for this scenario is shown in Figure 6.5.

For the evaluation, the Direct and Relaying schemes use the maximum transmit power (100mW) for all transmissions, whereas the SimTX uses the maximum transmit power (100mW) for the AP-r transmissions and variable transmit power for the r-d transmissions.

Since the SimTX scheme is intended to improve relaying performance, we are not very interested in the cases where direct transmissions are always

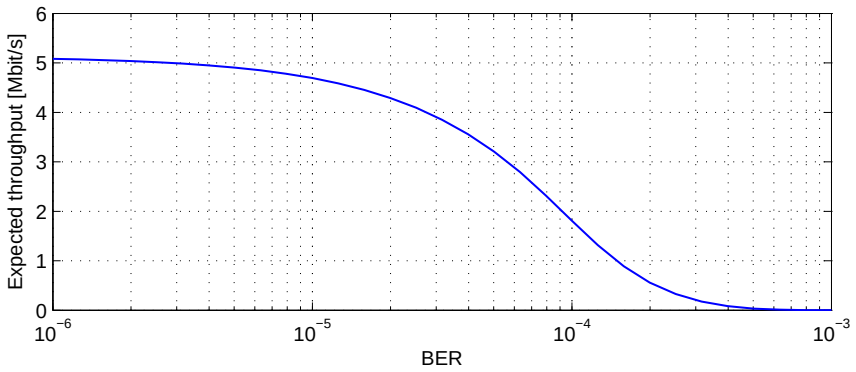


Figure 6.5: Expected average throughput for Rmax=7 and $B_{\mathrm{MSDU}} = 1024$ bytes.

superior. Therefore we pick the primary destination nodes randomly from nodes that are at least $30m$ from the AP. Figure 6.6 shows the throughput for each algorithm for different simulation runs. For each run, which corresponds to a different node layout, the throughput is calculated for an attempt to deliver a pair of packets to a primary and secondary destination node, respectively. The throughput calculation is done using the equations in section 6.4. We observe that the algorithms have different maximum throughput levels around 5, 2.6, and 3.5 Mbit/s, respectively. The Direct throughput fluctuates a lot, since some destination nodes cannot be reached successfully in one hop. The Relaying throughput is quite steady around half of the Direct, since it uses two consecutive transmissions. Finally, the SimTX algorithm improves the relaying performance thanks to the simultaneous transmissions.

Additionally, results of the embedded simulations for the SimTX algorithm (as described in section 6.5) are shown to evaluate the impact of the assumptions in equation (6.5). Specifically, in the simulation, the SINR is different for data and ACK since the ACK is transmitted the opposite direction, and dependent on if the other node transmits. The results show that the model is pessimistic in some cases as expected, especially when the throughput is between 1 and 2 Mbit/s. As this happens quite rarely, and Relaying performs better in most of these cases, we can use the model for parameter selection.

Figure 6.7 shows the power level selections by the SimTX algorithm that lead to the highest throughput, in the cases where the SimTX algorithm performed best. Here we can notice that the algorithm does take advantage of the possibility to adapt the transmit power.

Considering now the spatial layout of the considered scenario, Figure 6.8 shows us the positions at which each algorithm is preferred. Starting from the

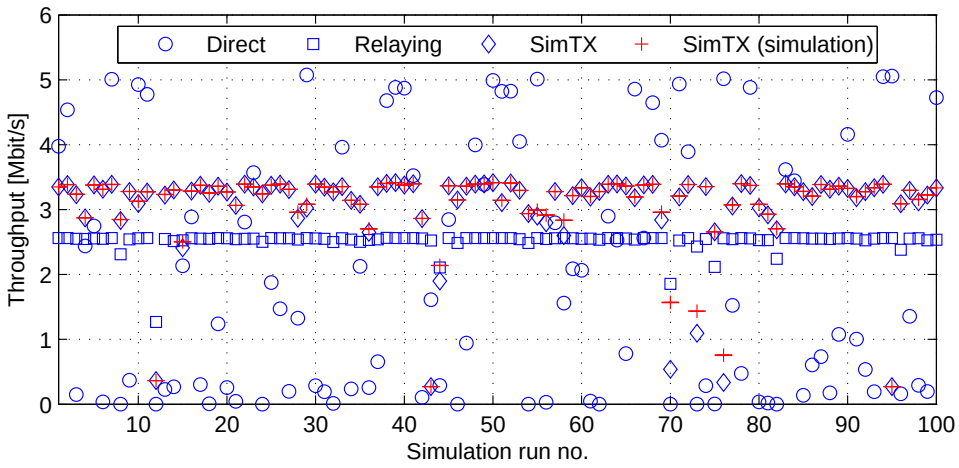


Figure 6.6: Throughput for the considered algorithms and comparison of model output and simulation of the 802.11 back-off algorithm in the case of two simultaneous transmissions.

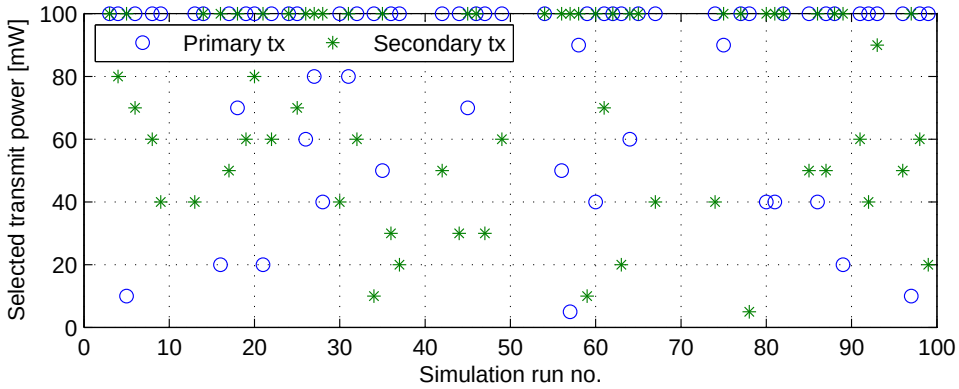


Figure 6.7: Transmit power selection by SimTX algorithm when SimTX is the preferred algorithm. The available power levels are shown in Table 6.1.

positions nearest to the AP, we see that the Direct algorithm is preferred at the positions that are within approximately 35 meters of the AP, since it is able to achieve a high throughput. Beyond this distance, both the Relaying and SimTX algorithms are represented, and they seem to be able to extend the range of the AP to achieve a reasonable throughput at any position in the $100x100m^2$ map.

A different view of this picture is given in Figure 6.9, where we see the density of destination nodes for each of the three algorithms. Here we also clearly see the how the Direct algorithm is preferred close to the AP in $(0, 0)$, but in this figure it is easy to see that SimTX is preferred in a ring at a larger

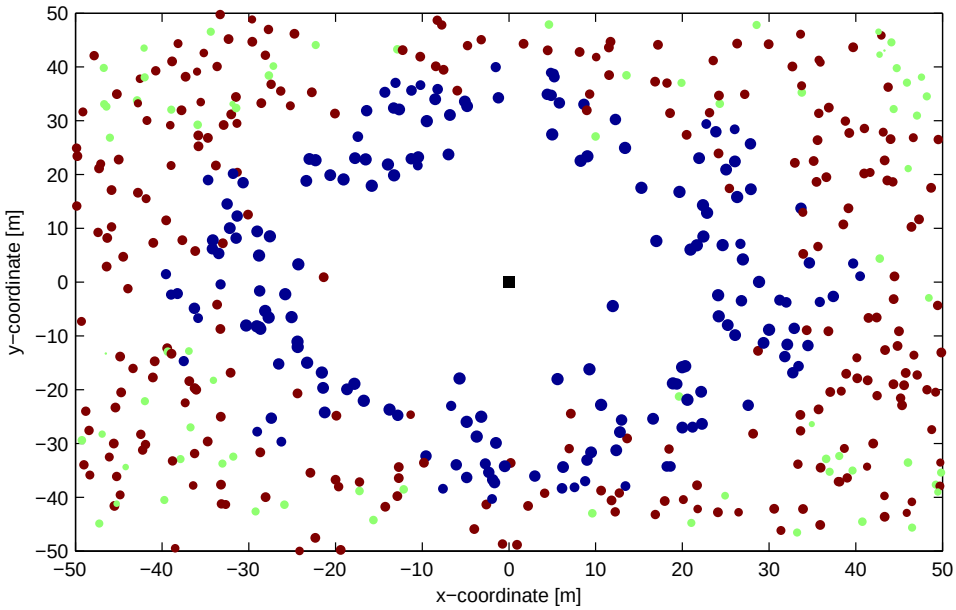


Figure 6.8: Points show destination node positions, where the color indicates the preferred algorithm (blue: Direct, green: Relaying, red: SimTX), and the size of the point is proportional to the throughput. Overlayed results from 250 runs.

distance than the Direct whereas the Relaying algorithm is mainly preferred in the corners.

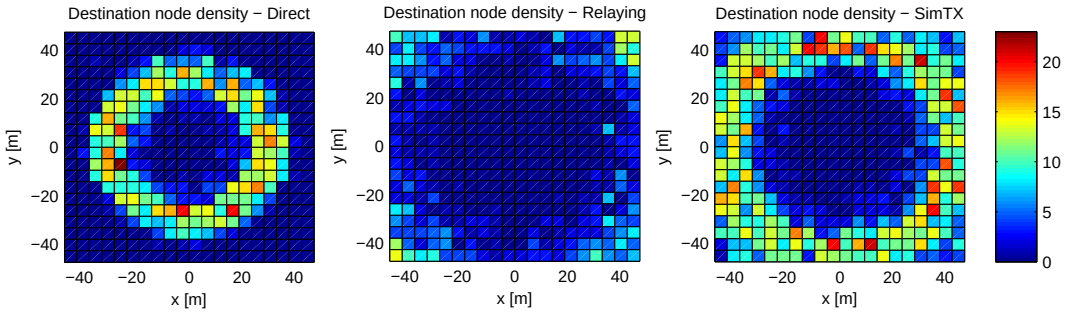


Figure 6.9: Density of destination node positions for each algorithm, based on 2500 runs.

Figure 6.10 shows a summary of the achieved throughput, when considering the cases where either Relaying or SimTX are the preferred algorithms. We see that the SimTX throughput is appr. 14% (2.45 to 2.8 Mbit/s) higher than for Relaying. If we consider the improvement from using only Relaying to using

the best of Relaying and SimTX in each case, an increase of approximately 20% (2.45 to 2.95 Mbit/s) is achieved.

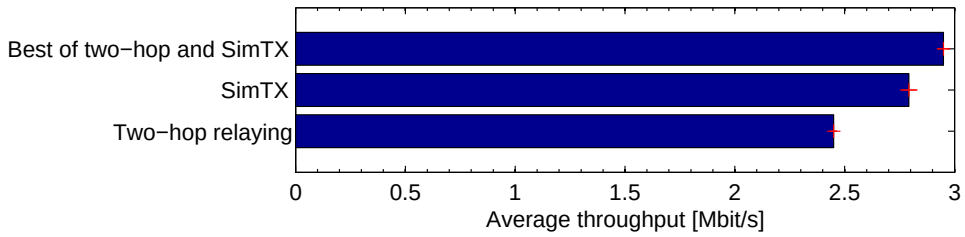


Figure 6.10: Comparison of the average throughput, for cases where relaying or SimTX is preferred, including 95% confidence interval. Based on 2500 repetitions.

So far, the analysis has assumed perfect knowledge of position information. Since position information is typically estimated using a localization system, which is not able to perfectly estimate the position, also the impact of inaccurate position information has been evaluated. For this, we have assumed that the position estimate of each mobile user is subject to a zero-mean two-dimensional Gaussian error. Figure 6.11 shows how the average throughput is affected when this localization error term is increased.

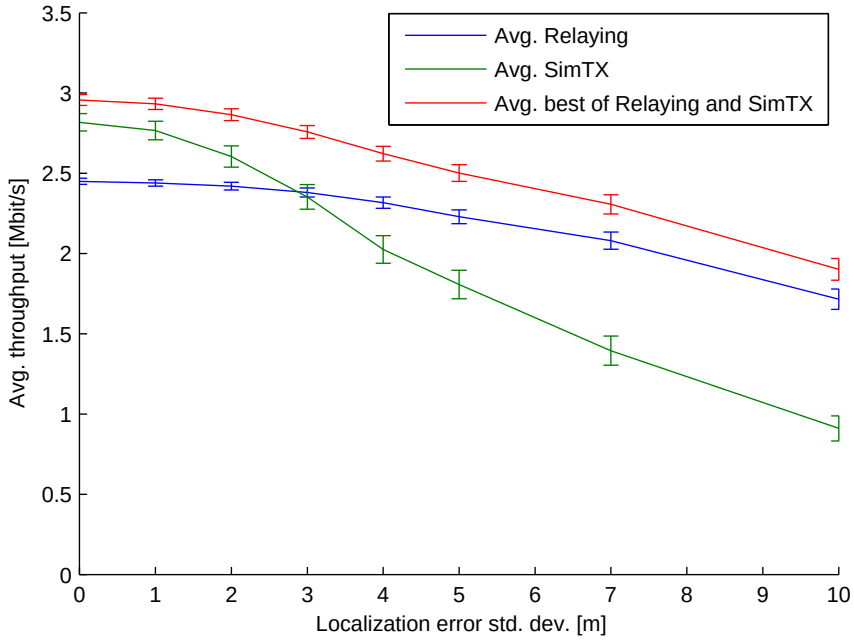


Figure 6.11: Performance impact of increasing localization error. Showing 95% confidence intervals.

Interestingly, the impact of the increasing localization error is different for the SimTX and Relaying algorithms. The SimTX algorithm clearly suffers more from the increasing error than the Relaying algorithm, as the achieved throughput decreases below that of the Relaying algorithm. Further, the improvement that the SimTX algorithm brings is decreasing, which is seen by the decreasing difference between the *Avg. Relay* and *Avg. best of Relaying and SimTX* curves in the figure. Considering that the SimTX algorithm depends on position information not only for relay selection like the Relaying algorithm, but also for the selection of the secondary destination node as well transmit power level of the relays, it makes sense that the impact of inaccurate information is bigger.

Finally, since the considered scenario does not include any kinds of cross-traffic or interference from neighboring WiFi hotspots, we would expect lower performance in a more realistic setting where this is included. It would be reasonable to expect that the SimTX algorithm is more sensitive to such cross-traffic and interference than the Relaying algorithm, since the Relaying algorithm is always transmitting at the highest possible transmit power, whereas the SimTX reduces the transmit power to limit the mutual interference between the simultaneous transmissions. But to which degree the performance is impacted, is hard to foresee and therefore an item for future work.

6.7 Conclusion

In this work we have proposed the SimTX algorithm that jointly optimizes the choice of relays and relay transmit power for two simultaneous relay-to-destination transmissions. For relay choice and transmit power selection, we have applied a model to calculate the expected MAC layer throughput when taking into account the BER, maximum limit of retransmissions and interference in case of simultaneous transmissions. The model allows the relaying scheme to choose the expectedly best relays and transmit power levels online.

Our results show throughput improvements of appr. 20% for the considered scenario compared to typical two-hop relaying in ideal settings with perfect location information. That is, we have shown that two simultaneous relay-to-destination transmissions can be beneficial despite the cross-interference they induce on each other. Further, we have shown that inaccurate location information has a bigger impact on the SimTX algorithm than on the two-hop Relaying algorithm, as the SimTX algorithm also uses the location information for selection of the secondary destination node as well as power adaptation.

In the previous chapter 5 the impact of mobility, delay, and inaccurate input information on the location based two-hop relaying scheme was analyzed. To a large extent, the tendencies of these results also apply to the SimTX algorithm. However, as the SimTX scheme was shown to suffer more from inaccurate location information than the two-hop relaying scheme, it is expected that it would also be more sensitive towards other types of errors than the two-hop relaying scheme.

In Chapter 4 the localization error standard deviation was found to be 2-3 meters. The results for the scheme showed that location errors in the order of 2-3 meters standard deviation, means that the scheme is still superior or on par with the two-hop relaying scheme.

During evaluation of the proposed algorithm, it was assumed that all data transmissions used the same payload size. This assumption may not always hold in practice, however, in many cases when the transferred data is part of a large file download or streaming media content, the payload is typically fragmented according to the Maximum Transmission Unit (MTU) size of the network. In principle the SimTX scheme could easily handle differently sized frames, the impact of this would just be a slightly lower throughput gain.

Further, this analysis has assumed that a fixed bit rate of 6 Mbit/s has been used for all nodes. Since a higher bit rate can be used when the signal quality is good, an obvious extension to the scheme would be to consider also bit rate adaptation, as in the related work [Chevillat et al., 2005, Hu and Tham, 2010, Liu et al., 2007]. Letting the SimTX algorithm consider also the different available bit rates in the optimization, would add another dimension to the optimization problem, which corresponds to the number of bit rates. With an increasing number of states to explore in order to find the optimal configuration, it would be beneficial to investigate approaches that are more efficient than the brute force state space search that the algorithm is currently using. One example would be the Partially Observable Markov Decision Process (POMDP) formulation that the authors propose in [Hu and Tham, 2010].

In the mentioned related work, most approaches are using link quality measurements in order to coordinate transmissions. Therefore, an interesting thought is whether the proposed scheme could have been realized without the use of location information. A central part of the SimTX algorithm is to be able to predict the mutual interference of simultaneous transmissions. This calculation would not be immediately possible without location information. The algorithms in [Hu and Tham, 2010, Chevillat et al., 2005, Liu et al., 2007] use past measurements and overheard transmissions in order to coordinate

transmissions. As such approaches need to learn the link qualities over time, they do not cope as well with node mobility, as a location based system.

References

Pierre Chevillat, Jens Jelitto, and Hong Linh Truong. Dynamic data rate and transmit power adjustment in ieee 802.11 wireless lans. *International Journal of Wireless Information Networks*, 12:123–145, 2005. ISSN 1068-9605. doi: 10.1007/s10776-005-0006-x.

T. ElBatt and A. Ephremides. Joint scheduling and power control for wireless ad hoc networks. *Wireless Communications, IEEE Transactions on*, 3(1): 74 – 85, 2004. ISSN 1536-1276. doi: 10.1109/TWC.2003.819032.

Zhengqing Hu and Chen-Khong Tham. Ccmac: Coordinated cooperative mac for wireless lans. *Computer Networks*, 54(4):618 – 630, 2010. ISSN 1389-1286. doi: 10.1016/j.comnet.2010.02.001. Advances in Wireless and Mobile Networks.

IEEE. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999)*, pages C1–1184, 12 2007.

T.S. Kim, H. Lim, and J.C. Hou. Improving spatial reuse through tuning transmit power, carrier sense threshold, and data rate in multihop wireless networks. In *Proc. of the 12th annual int'l conference on Mobile computing and networking*, page 377. ACM, 2006.

W. Li, P. Lv, Y. Chen, and M. Xu. POCOSIM: a power control and scheduling scheme in multi-rate wireless mesh networks. *Ubiquitous Intelligence and Computing*, pages 474–488, 2010.

H. Lim, C. Lim, and J.C. Hou. A coordinate-based approach for exploiting temporal-spatial diversity in wireless mesh networks. In *Proc. of the 12th annual int'l conference on Mobile computing and networking*, page 25. ACM, 2006.

P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S.S. Panwar. CoopMAC: A cooperative MAC for wireless LANs. *IEEE Journal on Selected Areas in Communications*, 25(2):340, 2007.

Hui Ma, Jing Zhu, Sumit Roy, and Soo Young Shin. Joint transmit power and physical carrier sensing adaptation based on loss differentiation for high density ieee 802.11 wlan. *Computer Networks*, 52(9):1703 – 1720, 2008. ISSN 1389-1286. doi: DOI:10.1016/j.comnet.2008.02.008.

T. Nadeem, L. Ji, A. Agrawala, and J. Agre. Location enhancement to IEEE 802.11 DCF. In *Proc. IEEE INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 1, 2005.

Daji Qiao, Sunghyun Choi, Amit Jain, and K.G. Shin. Adaptive transmit power control in ieee 802.11a wireless lans. In *Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE Semiannual*, volume 1, pages 433 – 437 vol.1, 2003. doi: 10.1109/VETECS.2003.1207577.

7

Location based Handover Optimization

The contribution presented in this chapter concerns handover between WiFi and cellular networks in large scale heterogeneous scenarios. Specifically, we investigate how the knowledge of location information can help in the multi-system handover (HO) decision. The main problem we address is how location information can be used to guide a mobile device's selection between the ubiquitous cellular network and any locally available WiFi networks. As stated in chapter 1, a main assumption in this work is the availability of a database that contains the average throughput of all available networks at any position for the considered geographical area.

7.1 Introduction

As the amount of mobile data is expected to grow appr. 108%/year until 2014 according to [Cisco Systems, 2010], the cellular networks may not be able to satisfy this demand alone. The motivation for this work is to develop methods for offloading the network operators cellular networks using Wi-Fi networks when this can improve the mobile user's connectivity in terms of e.g. throughput. In the two experimental performance comparisons of 3G and Wi-Fi in [Gass and Diot, 2010] and [Deshpande et al., 2010], it is clear that if a terminal is able to use the network that offers the highest throughput at any given time, an overall performance increase is possible, while the cellular network is offloaded when Wi-Fi networks are used. In order to make use of the Wi-Fi the user terminal needs to make a handover from the cellular network to the Wi-Fi access points along the user's movement trajectory, as sketched in the scenario overview shown in Figure 7.1.

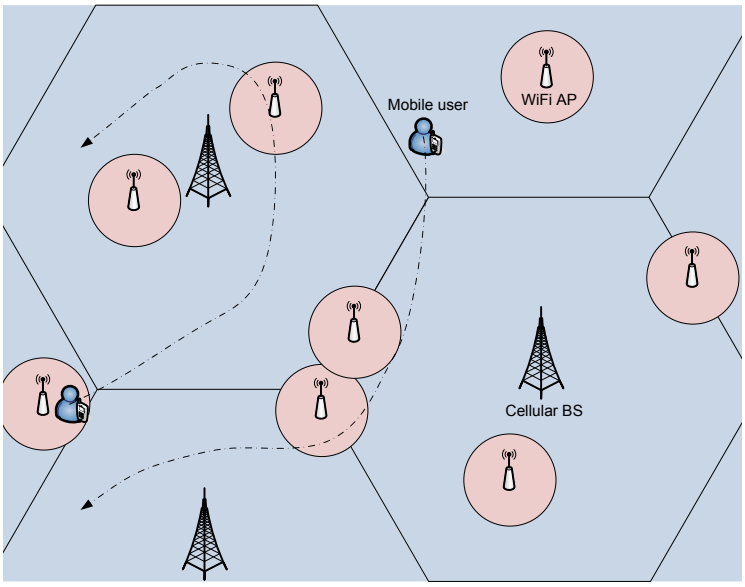


Figure 7.1: Large scale scenario with ubiquitous cellular coverage and scattered Wi-Fi access points. The two mobile users may potentially achieve a higher throughput if they handover to the Wi-Fi networks, when in range. The dash-dotted lines indicate possible movement trajectories. Notice: copy of Figure 1.1.

As discussed in section 1.3 on page 8, the work presented in this thesis assumes that the Media Independent Handover (MIH) framework is used for coordination and execution of handovers. The MIH framework does however not specify how handovers are triggered and to which network the terminal

should connect. The contributions presented in this chapter fill in this gap, by considering specific location based algorithms for the handover trigger or handover decision. Also, the MIH framework is assumed to be used for enabling access to a MIH Information Server (MIIS), which is used for storing environment fingerprints, which can be used to estimate the achievable throughput, and location information of network entities. Such a database of environment fingerprints is necessary for performing localization through fingerprinting, and it may therefore already be available.

Additionally, it is assumed that Mobile IP (MIP) is used to avoid that the IP address changes when the mobile device associates with a new network, thus allowing service continuity for applications, meaning that the considered handover algorithms need only to determine which network(s) to handover to, and when.

In general, handovers are usually referred to as horizontal handover (HHO) or vertical handover (VHO). The HHO refers to a handover between access points or base stations using the same access technology, whereas the VHO refers to handovers between networks using two different access technologies, e.g., from IEEE 802.11a to a 3G cellular network. The VHO always changes the IP configuration, since the VHO requires that the data link and physical layers of the protocol stack are switched to the new access technology. The HHO does not always cause the IP configuration to be changed, for example in the case where a user is moving between two APs belonging to the same Basic Service Set (BSS) or Extended Service Set (ESS), the handover is handled transparently on the data link layer. However, when moving between BSSs or ESSs, the IP configuration cannot be kept.

As this chapter focuses on exploiting available Wi-Fi networks, the analysis focuses on VHOs between the cellular network and Wi-Fi networks as well as HHOs between individual Wi-Fi networks, which are assumed not to belong to the same BSS or ESS. In both cases the IP configuration of the mobile station changes, but this can be handled through network layer functions such as for example MIP as described in chapter 1. For simplicity, both types of handovers are in the following referred to commonly as handovers.

Handovers are typically categorized as either hard (also called break-before-make) or soft (also called make-before-break). The hard handover is needed in cases where a mobile terminal only has a single radio, which it uses to establish connections. When switching between networks, the terminal therefore must break the connection to a network before it can establish a connection to the new network. The soft handover can be used by terminals that have more than one radio, and are therefore able to configure the second interface for a seamless

handover using MIP. For handovers between cellular and Wi-Fi networks, the soft make-before-break handover can be used for mobile devices that have a dedicated radio for cellular connectivity and a dedicated radio for Wi-Fi connectivity. However, in the cases where a handover is performed between two Wi-Fi networks, it is necessary to use the hard break-before-make approach, since typically, only one Wi-Fi capable radio is available, and the Wi-Fi device can only associate with one network at a time. In [Ramachandran et al., 2006] the authors propose different handover algorithms that use either one or two Wi-Fi adapters and different strategies for probing available networks. A seamless handover is possible with some modifications to the protocol stack, but only in a two-card setup. The authors aurgue the handover time of a hard handover with an off the shelf IEEE 802.11 adapter, can be in the order of a second, when the terminal needs to probe all channels to find a suitable network.

Every handover that a mobile device performs from one network to another has a cost associated to it. This cost may be mainly in the form of "lost" throughput while the handover is executed if it is a hard handover or in the form of signaling overhead for the AP association, DHCP look-up, IP configuration, as well as reconfiguration of the MIH and MIP functionalities if it is a soft handover. In either case, it is preferable that the amount of handovers is limited. In the present contribution, for simplicity, all considered handovers (cellular to Wi-Fi, Wi-Fi to cellular, and Wi-Fi to Wi-Fi) have been assumed to be of the break-before-make type. Thereby, the cost of a soft handover will also be measured in lost throughput, however this simplification of the handover cost is justified since it will help to keep the number of handovers low.

Given the cost of each handover and the possibility of having many candidate networks in dense urban environments, the timing and choice of target network for each handover is crucial for mobile users. An important functionality for mobile devices intended for use in heterogenous network scenarios is therefore the HO algorithm, which should make the best selection of target network and handover time. The goodness of a handover can be given in terms of different metrics such as: application throughput, application jitter/delay, or signaling overhead. Some users may be more interested in maximizing the throughput (e.g., for downloading or streaming), while others may value low delay and jitter (e.g., for VoIP calls). From the operator's viewpoint a low number of handovers (to keep signaling overhead low) may be preferred.

7.2 Related Work

In the literature many HO decision algorithms already exist. The references [Yan et al., 2010] and [Kassar et al., 2008] present two recent surveys on vertical HO decision algorithms for heterogeneous wireless networks. In the following, an overview is given over some of the works found herein in addition to a few more examples of recent works.

HO decision algorithms can be categorized in many different ways, with some examples of categories being:

Input data such as RSS values, location information, and bandwidth.

HO criteria such as RSS threshold, connection time, and power consumption.

HO algorithm type such as threshold/hysteresis, optimization, and cost function.

Decision type such as reactive or predictive.

In the following overview, the existing work is categorized according to the HO criteria, similar to the categorization in [Yan et al., 2010], however with a special emphasis on the use of location information in the algorithms.

RSS based algorithms include the algorithms that use measured RSS information as the main HO decision parameter. Because of the widespread availability of RSS information in most wireless equipment, a comparison of RSS levels has traditionally been used as the handover trigger. The survey [Yan et al., 2010] mentions a number of such basic schemes.

The two RSS based schemes proposed in [Zahran et al., 2006] and [Yan et al., 2008] both use only RSS information of cellular and WLAN networks to predict the time within coverage of candidate networks. By estimating the rate of change of the measured RSS values, the remaining time within coverage can be estimated. This predicted time is used to decide when to handover to a given network and in which cases a handover is not beneficial. In both papers the prediction of RSS values is only used for timing the handover. The network to handover to is selected based on instantaneous RSS levels only.

In [Mohanty and Akyildiz, 2006] a related RSS based algorithm for next generation wireless systems is presented, in which a dynamic RSS threshold is

calculated based on different scenario parameters such as path loss exponent, cell size, user movement speed, and handover latency. This threshold is used to ensure that the handover is executed before the user has moved out of coverage.

Bandwidth based algorithms are using the available bandwidth for different networks as the handover criteria. Compared to RSS based algorithms, the choice of bandwidth based algorithms can more easily be conditioned on certain Quality of Service (QoS) requirements, which makes it easier to accept or reject candidate networks.

An example of a bandwidth based scheme is described in [Yang et al., 2007] and is targeting handover between WLAN and Wideband Code Division Multiple Access (WCDMA) systems. The algorithm will handover and stay at the network that has the highest SINR, which directly determines the achievable bit rate. The SINR of the WLAN is converted so that it is directly comparable to the WCDMA SINR. Obviously, there is a risk of many unnecessary handovers of such a reactive scheme if the SINR levels of two or more networks are close and varying. This could lead to frequent handovers, resulting in only a marginal (if at all) overall gain in achieved performance.

The handover decision scheme in [Lee et al., 2005] is intended for WLAN and Wide-Area Access Network (WAAN) scenarios. The scheme estimates the available bandwidth in WLAN via the QBSS field in IEEE 802.11e beacons ([IEEE, 2007]), which gives the number of stations, channel utilization, and packet loss rate. Per default, the scheme connects to the WAAN, but it continuously monitors RSS of beacons from nearby WLANs and makes handover if RSS has been sufficient for a certain time and if the QoS bandwidth requirements can be fulfilled by the considered WLAN. The analysis only considers a single WLAN and a single WAAN network, which makes it difficult to judge how the algorithm would cope with multiple possible WLANs.

Cost function based algorithms are combining different input parameters into a single cost function, and the result of this cost function is then used as handover criteria.

One example of a cost function based algorithm is given in [Chen and Shu, 2005], where a Location Server Entity provides information such as coverage area, bandwidth and latency of nearby wireless networks, which is combined to express the QoS of the current and available networks. The mobile terminal uses this QoS criteria for HO network selection.

The authors of both papers [Sun et al., 2008] and [Zhang et al., 2006] have formulated the HO decision as Markov decision processes, taking into account different parameters such as: connection duration, QoS parameters, location information and predicted movements, network access cost, and the signaling load incurred on the network. The proposed algorithms are shown to work well compared to state of the art algorithms. However, in both papers the authors assume that the achievable throughput of the available networks are constant within the coverage region, which is not the case in practice as shown in [Gass and Diot, 2010].

A context-aware vertical HO approach that can integrate a variety of wired and wireless technologies (2.5G, 3G, 4G, WLAN, and Bluetooth) is proposed in [Balasubramaniam and Indulska, 2004]. The first step in the HO process is to select a subset of candidate networks by using location information. In the second step an Analytic Hierarchy Process method is employed to maximize user device preferences and application bandwidth, while minimizing jitter, delay, and loss as well as bandwidth fluctuations, to optimize the user perceived QoS. Evaluation results are presented from a prototype implementation but are not compared to state of the art approaches.

Combination algorithms use a rich combination of input information, which is difficult to handle analytically. Instead, algorithms such as fuzzy logic or neural networks are applied.

In [Pahlavan et al., 2000] the authors propose a HO algorithm based on neural networks. The method is shown to be superior compared to simple RSS threshold and hysteresis-based schemes, however the algorithm requires substantial training beforehand in form of RSS traces and desired outputs for the different situations that may occur. In practice, it may be difficult to foresee all possible situations that can arise.

In [Xia et al., 2007] the authors present a fuzzy logic based HO decision algorithm, which uses current RSS, predicted RSS, and available bandwidth as input parameters. The fuzzy algorithm uses five quality levels for the RSS parameters and three quality levels for the bandwidth parameter. For each candidate network a performance evaluation value (PEV), which is a fuzzy combination of the three parameters' quality levels, is calculated, and the target network is the network with the highest PEV. The performance of the approach is however not shown or compared to state of the art approaches.

Based on the reviewed literature, it is clear that a handover decision algorithm should seek to limit the number of unnecessary handovers while satisfying the user's QoS. In the present contribution, we propose two HO decision algorithms that work proactively by using movement prediction to predict the expected future connection quality of available networks and plan the HO so that QoS in terms of throughput is maximized and the number of handovers is kept at a minimum. The algorithms use a distance dependent throughput model to predict the achievable throughput of each network. This is quite similar to the work of [Sun et al., 2008] and [Zhang et al., 2006], but in both papers the authors assume constant throughput within the coverage region for the MDP problem formulations, whereas our optimization algorithm uses differentiable functions that represent the predicted throughput to determine the optimal handover decisions.

Initially, only a single user is considered, and thus the load of other users is not taken into account. An obvious extension would be to consider the available resources using the 802.11e based approach described in [Lee et al., 2005].

7.3 System Model

We consider a single user terminal that can connect to either a cellular network work or a Wi-Fi network. As discussed in section 7.1, it is assumed in the following that handovers are of the break-before-make type. The terminal has the location $X(t) = [x(t), y(t)]^T$ at time t. This location is however not known, but it is estimated as $\hat{X}$ using a localization system, which can be GPS or network-based as described in chapter 4. The mobile terminal's past movement trajectory given by $\hat{X}(t \leq t_\mathrm{c})$, where t_c is the current time, is based on interpolation of previous location estimates. The mobile terminal's predicted future movement trajectory $\hat{X}^*(t_\mathrm{c} \leq t \leq t_\mathrm{c} + W)$, where W is a time window specifying the prediction horizon, is based on an extrapolation of the past trajectory, t_h seconds back in time. In the following we denote W as the *look-ahead window*. See Figure 7.2 for an illustration.

By connecting to an access point or a base station with index a, whose coordinates are known a priori, the mobile terminal at location $\hat{X}$, at time t, achieves a throughput Ω:

$$\Omega_a(t) = S_a(X(t)) + V \tag{7.1}$$

where the random variable V accounts for variations in the actual throughput, caused by non-deterministic factors such as small and medium-scale fading.

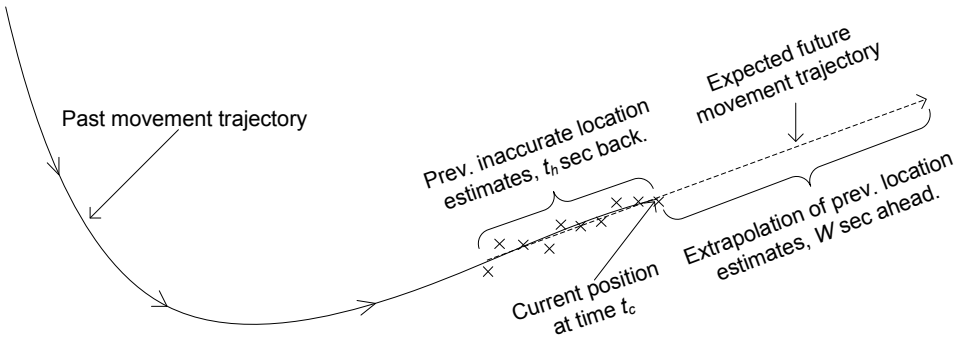


Figure 7.2: The predicted future movement trajectory within the time window W is created as an extrapolation of previous location estimates.

We assume that V is a zero-mean gaussian stochastic variable, characterized by the standard deviation σ_{TP}.

The expected throughput for a network a is achieved from the database as the approximation:

$$\Omega_a^*(t) = S_a(\hat{X}^*(t)) \tag{7.2}$$

where $S_a(\hat{X}^*(t))$ is the expected throughput at the predicted location $\hat{X}^*(t)$ at time t, given the path-loss of the link between the terminal and the AP or BS.

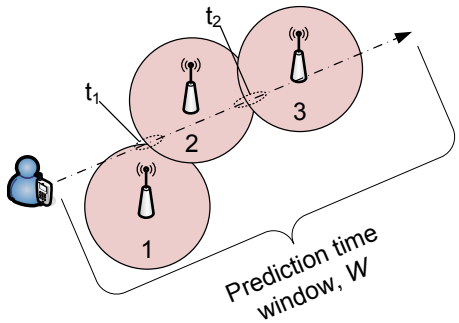
The mobile terminal can choose to connect to a different AP/BS by performing a handover (HO). Within the time window $[t_c; t_c + W]$, the terminal may perform a sequence of HOs $\mathcal{H}$, defined as:

$$\begin{aligned} \mathcal{H}^K = &\{(a_i, t_i),\ i = 1 \ldots K\}, \\ &a_{i-1} \neq a_i, \\ &t_c < t_1 < t_2 < \ldots < t_K < t_{\mathrm{c}} + W \end{aligned} \tag{7.3}$$

which describes a HO to network a_i at time t_i, where index K denotes the number of HOs in the sequence. Notice that the target network in a handover is never the same as the source network.

Figure 7.3 shows an example scenario and a corresponding handover sequence $\mathcal{H}$ for $K = 2$.

Each performed HO in the sequence $\mathcal{H}$, may incur a cost due to lost connectivity while switching from one network to another or signaling overhead, since this switch requires steps such as AP association, Dynamic Host Configuration Protocol (DHCP) look-up, and Internet Protocol (IP) address configuration.



Handover sequence $\mathcal{H}$ for $K = 2$:	
Network	**Time**
$a_1 = 2$	t_1
$a_2 = 3$	t_2

Figure 7.3: Example scenario and handover sequence. Since the mobile user barely comes within reach of network 1, the networks 2 and 3 are probably the preferable networks in this example. The dotted regions indicate the likely places of the two handovers at time t_1 and t_2.

In this work we assume that this cost denoted C_a is a downtime or handover delay, where the throughput is zero. C_a depends only on the target network. The actual handover cost will in practice also include a certain signaling overhead, which motivates to keep the number of handovers low. As the achieved throughput is zero for the duration of each handover, optimizing for the highest throughput should also result in a low number of handovers for not too small values of C_a.

In order to experience the best performance in terms of throughput, the mobile terminal needs to determine when is the best time to perform HO(s) and which network(s) to connect to, taking into account the handover cost C_a.

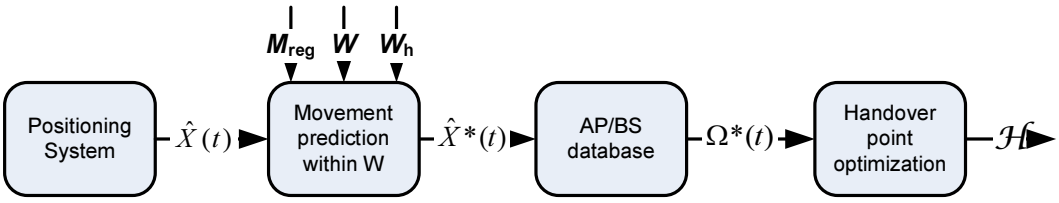


Figure 7.4: Look-ahead prediction algorithm principle. M_{reg} is the order of regression, W is the look-ahead window length, W_{h} is the look-back window length.

7.4 Look-ahead Prediction Algorithms

In the following, we will describe the considered HO decision algorithms. The first two algorithms are so-called look-ahead prediction algorithms that determine a HO sequence $\mathcal{H}$, as depicted in Fig. 7.4. The main assumption for these algorithms is that a fingerprinting database that contains the average throughput of all available networks for the considered geographical area is available.

7.4.1 Optimal K-Handover Look-ahead Algorithm

The optimal sequence of HOs that maximizes the throughput within the time window W, with exactly K HOs, may be defined as:

$$\mathcal{H}^K_{\text{opt}} = \underset{\mathcal{H}^K}{\text{argmax}}(f(\mathcal{H}^K)) \tag{7.4}$$

$$f(\mathcal{H}^K) = \int_{t_0}^{t_1} \Omega_{a_0}(t)dt + \sum_{i=1}^{K} \left(\int_{t_i+C_{a_i}}^{t_{i+1}} \Omega_{a_i}(t)dt \right) \tag{7.5}$$

where $t_{K+1} = t_0 + W$. The integration of $\Omega_a(t)$ over time corresponds to the throughput experienced when connected to network a along the predicted movement trajectory, so $f(\cdot)$ is the total throughput achieved within W for a given HO sequence. Notice that we determine the optimal $\mathcal{H}^K_{\text{opt}}$ for each of the cases $K = \{1, 2, \dots K_{max}\}$ separately, and then select the best number of HOs:

$$K_{opt} = \underset{K}{\text{argmax}}(f(\mathcal{H}^K_{\text{opt}})),\ K = 1 \dots K_{\text{max}}. \tag{7.6}$$

This is solved by iterating over all considered values of K and selecting K_{opt} as the K that leads to the highest throughput.

In order to determine the optimal HO sequence for a value of K, we consider all possible combinations of networks to handover from and to, as well as the candidate HO times, which will be defined subsequently.

The N possible network combinations are:

$$\begin{aligned} A_n =& \{a_0^n\ a_1^n \cdots a_i^n \cdots a_K^n\},\ n = 1 \dots N, \\ & a_{i-1}^n \neq a_i^n, \\ & a_0^n = a_0^j, j = 1 \dots N. \end{aligned} \tag{7.7}$$

For every HO i between two consecutive networks in A_n, there is a set of M time instants that are candidates for optimal HO points between these two networks. This set is defined as:

$$T_n^i = \{t_i^{n,1}\ t_i^{n,2} \cdots t_i^{n,m} \cdots t_i^{n,M}\}. \tag{7.8}$$

Each unique combination of networks and time instants from A_n and $T_{n,m}$ constitute a unique sequence:

$$\mathcal{H}_{n,m}^K = \{(a_i^n, t_i^{n,m}),\ i = 1 \dots K\}. \tag{7.9}$$

Now, given the n'th combination of networks A_n we can determine the set of candidates for optimal handover points T_n^i, as the t_i's that satisfy:

$$\frac{df(\mathcal{H}_n^K)}{dt_i} = 0 \tag{7.10}$$

since these points result in either maxima or minima for $f(\mathcal{H}_n^K)$, which expresses the total throughput in W.

If we define $O_a(t)$ as the primitive function of $\Omega_a(t)$, that is $\Omega_a(t) = \frac{dO_a(t)}{dt}$, we can rewrite $f(\mathcal{H}_n^K)$ as:

$$\begin{aligned} f(\mathcal{H}_n^K) &= O_{a_0^n}(t)\Big|_{t_0}^{t_1} + \sum_{i=1}^{K} \left(O_{a_i^n}(t)\Big|_{t_i+C_{a_i}}^{t_{i+1}} \right) && (7.11) \\ &= O_{a_0^n}(t_1) - O_{a_0^n}(t_0) \\ &\quad + \sum_{i=1}^{K} \left(O_{a_i^n}(t_{i+1}) - O_{a_i^n}(t_i + C_{a_i}) \right). && (7.12) \end{aligned}$$

Such optimization by differentiation of course requires that the primitive function of $O_a(t)$ is continuously differentiable.

Differentiation of $f(\mathcal{H}_n^K)$ with respect to t_i, reduces to:

$$\frac{df(\mathcal{H}_n^K)}{dt_i} = \Omega_{a_{i-1}}(t_i) - \Omega_{a_i}(t_i + C_{a_i}). \tag{7.13}$$

Setting this expression equal to zero and finding all solutions for every t_i gives the candidate handover points T_n for the n'th combination of networks. In cases with multiple solutions for each t_i, the algorithm tries out all found solutions.

In short, the complete algorithm can be described as shown in Algorithm 1, where the function `generate_combinations`(N_a, K, a_0) generates all possible combinations of networks as specified in eq. (7.7). N_a is the number of networks available within W, and K is the number of allowed HOs with W. Further, N is the number of network combinations in A.

for $K = 1 \ldots K_{max}$ **do**
  $A =$ `generate_combinations`(N_a, K, a_0)
  for $n = 1 \ldots N,\ N = |A|$ **do**
    $\mathcal{H}_n^K = A_n$
    for $i = 1 \ldots K$ **do**
      Solve $\frac{df(t,\mathcal{H}_n^K)}{dt_i} = 0 \ \rightarrow \ T_n^i$
    end
    for $m = 1 \ldots M$ **do**
      $\mathcal{H}_n^K = \{A_n, T_{n,m}\}$
      $s(K, n, m) = f(t, \mathcal{H}_{n,m}^K)$
    end
  end
end
$(K, n, m) = \max\{s(K, n, m)\}$
$\mathcal{H}_\mathrm{opt} = \mathcal{H}_{n,m}^K$

Algorithm 1: Optimal K-HO algorithm.

The HO algorithm is run periodically, looking a time W ahead and looping over the K_max possible handovers to determine the optimal number of handovers as described in eq. (7.6). The HOs are done as planned, and after time W has passed since last run, the algorithm is run again. However, due to uncertainties caused by localization inaccuracy and unknown future movements, the predicted behaviour, which is used to calculate a HO sequence, is expected to become less trustworthy with increasing lengths of the look-ahead window.

Algorithm complexity The determining factor for the complexity of the algorithm is the number of different network combinations that the algorithm is trying out.

Considering the constraint in eq. (7.7): that a HO is always to a different network than the current, the number of entries in A_n becomes:

$$N = (N_\mathrm{a} - 1)^{K+1}. \tag{7.14}$$

From this it is clear, that for a large number of available networks, trying out all combinations can become infeasible as the number of HOs K increases.

In order to keep the complexity low, we consider only the networks whose expected throughput exceeds that of the cellular network within the window W. Hereby, the number of networks that the algorithm considers, N_{a}, is reduced.

7.4.2 Heuristic Look-ahead Algorithm

In addition to the previous algorithm that determines the optimal K-HO solution within the window W, we also consider a less complex heuristic look-ahead algorithm.

This algorithm always tries to handover to and stay on the network with the highest expected troughput at any given time. However, only if the gain of performing the handover exceeds the cellular network throughput by more than a fixed threshold ρ.

Assume that $\{t_1\ t_2\ \cdots\ t_j\ \cdots\ t_J\}$ where $t_j < t_{j+1}$ is a list of timestamps for when the network with the highest expected throughput, $a^{\max}$, changes. This can be determined by calculating the crossing points of all corresponding throughput functions. Then the preferred network for the j'th timespan $(t_j; t_{j+1}]$ is:

$$a_j^{\mathrm{pref}} = \begin{cases} a_j^{\max} & \text{if } \int_{t_j}^{t_{j+1}} (\Omega_{a_j^{\max}}(t) - \Omega_1(t))dt - \rho > 0 \\ 1 & \text{otherwise} \end{cases} \tag{7.15}$$

where $\Omega_1(t)$ is the throughput function of the cellular network and ρ is a threshold, used to filter out unhelpful HOs, which is set as defined in Table 7.2.

Finally, the heuristic sequence is:

$$\mathcal{H}_{\mathrm{heu}} = \{(a_j^{\mathrm{pref}}, t_j),\ t = 1 \ldots J\} \tag{7.16}$$

Notice that for simplicity, this heuristic algorithm does not take the HO delay into account when calculating the timestamps in the HO sequence.

7.5 Reference Algorithms

For comparison, we include the following two reference algorithms:

7.5.1 Maximum Throughput Algorithm (C=0)

This algorithm outputs the maximum instantaneous throughput of all available network. This corresponds to always performing a handover to the network with the highest throughput, under the assumption of no handover delay, i.e., $C = 0$. This algorithm should not be seen as a practical algorithm, but it only serves as an upper bound on performance. Notice that for practical systems where $C > 0$, the bound is not tight.

7.5.2 Hysteresis based HO Algorithm

This algorithm triggers a HO to another network, if the instantaneous throughput of another network exceeds the instantaneous throughput of the currently connected network by more than a threshold β_{hyst}. In a practical system the instantaneous throughput would be calculated from the instantaneous SNR, or the threshold would be given as an SNR-threshold. If more than one other network exceeds the threshold, the network offering the maximum throughput is chosen. That is, a HO is initiated if the set of candidate networks A_{hyst} is not empty. A_{hyst} consists of the networks a that fulfill:

$$\Omega_a(t) > \Omega_{a_0}(t) + \beta_{\text{hyst}} \tag{7.17}$$

where a_0 is the currently connected network.

The network to handover to is selected as:

$$a_{\max} = \underset{a \in A_{\text{hyst}}}{\operatorname{argmax}}(\Omega_a(t)). \tag{7.18}$$

Contrary to the look-ahead prediction algorithms, this algorithm is a greedy algorithm that does not plan ahead, but decides when to handover based on the instantaneous throughput.

7.6 Evaluation Methodology

The proposed location based prediction algorithms have been evaluated and compared to the reference algorithms through an extensive simulation study, which has been implemented in matlab. The simulation results are based on an average of multiple simulation runs, where each of the simulation runs consists of the following steps:

1. Generate a random AP layout, where the nmber of APs specified in Table 7.2 are distributed randomly according to a uniform distribution, over the considered scenario area.

2. Generate a random movement trajectory from a randomly selected starting point (uniformly over whole area), with a step size of 1 second using the wrap around mobility model described in section 2.4.3, with the parameters and duration listed in Table 7.2.

3. Generate a vector for each network, which contains the instantaneous throughput for each step along the movement trajectory. The instantaneous throughput is calculated as the path loss dependent throughput given by the throughput model in Chapter 3 with an added zero-mean Gaussian random variable, defined by the standard deviation σ_{TP}.

4. For each point defining the movement trajectory, generate a corresponding set of location estimates, mimicking the output of a localization system. These location estimates are generated by adding a zero-mean Gaussian random variable, defined by its standard deviation σ_{pos}.

5. For each of the considered algorithms, step through the generated movement trajectory and run the handover decision algorithms at fixed intervals. The location based and non-location based algorithms are working slightly differently, as explained in the following paragraphs:

   The optimal K-HO and heuristic algorithms are executed with W seconds intervals. For each execution, first the movement prediction algorithm described in section 7.6.1 is executed. The inputs to the movement prediction algorithm are the location estimates for the previous W_{h} seconds. Now, along the predicted trajectory, the path loss dependent throughput in relation to each network, is calculated and fitted to polynomial functions as described in section 7.6.2. A separate throughput function is calculated for each network. These throughput functions are given as inputs to the handover decision algorithms in section 7.4, which each give as output a handover sequence $\mathcal{H}$. These handover sequences describe the handovers that each algorithm has decided upon for the W seconds time window. The simulation now jumps W seconds forward and executes the optimal K-HO and heuristic algorithms again. This is repeated until the specified duration of the simulation runs has elapsed. After the simulation run finishes, the handover sequences $\mathcal{H}$ for each prediction window are concatenated in order to obtain a complete handover sequence for the whole simulation run.

The reference algorithms defined in section 7.5 are both greedy algorithms that only consider the instantaneous throughput of the available networks for the handover decision. This means that the algorithms are executed at every time step. The output from each of these algorithms is also a handover sequence for the whole simulation run.

6. The final step of the simulation is the actual evaluation of the achieved throughput for each of the algorithms. This is done separately for each algorithm by summing up the instantaneous throughput along the movement trajectory for the selected networks specified by the handover sequence $\mathcal{H}$. Every time a handover is made, the achieved throughput is zero for C seconds, due to the handover delay. Notice that the *Maximum Throughput* reference algorithm is assumed to always have $C = 0$ and is therefore not subject to the handover delay when calculating the achieved throughput.

Finally, the average throughput and 95% confidence interval over all simulation runs are determined for each algorithm.

In the following two sections, the used movement prediction algorithm and the throughput functions are described in details.

7.6.1 Movement Prediction

For the look-ahead HO algorithms we use a linear movement prediction algorithm that uses historical location measurements within a time window W_h to predict a direction and constant speed of the mobile user, W seconds ahead. This prediction algorithm is based on Total Least Squares (TLS) regression and is described in details in the following. The movement prediction could also be done using a Kalman Filter, such as the one used for localization in Chapter 4. In the present contribution the localization is however not included, but represented abstractly through a localization error term.

First, the direction of movement is determined using a 1st order TLS regression. TLS is used, since it minimizes the perpendicular distance to the regression line and not only the vertical distance as ordinary LS does. The TLS is in this work realized using the Principal Component Analysis (PCA) method [Jolliffe, 2002], however the TLS solution can also be found using singular value decomposition or eigen vector decomposition, as discussed in [Simoncelli, 2003].

The PCA method gives a vector of unit length along which, the variance of the data is the highest. But the PCA does not tell if the movement direction is along or opposite the resulting vector, denoted $\hat{\mathbf{w}}$. Therefore we use the first and last historic data points to determine the sign of the direction vector as:

$$\hat{\mathbf{u}} = \begin{cases} \hat{\mathbf{w}} & \text{if } \text{sign}((\mathbf{a}^T\hat{\mathbf{w}})\hat{\mathbf{w}}) = \text{sign}(\hat{\mathbf{w}}) \\ -\hat{\mathbf{w}} & \text{otherwise} \end{cases} \tag{7.19}$$

where $\mathbf{a}$ is the last and first historic data points subtracted:

$$\mathbf{a} = \hat{X}(t_\text{c}) - \hat{X}(t_\text{c} - W_\text{h}) \tag{7.20}$$

where $\hat{X}(t_\text{c})$ is the estimated current location and $\hat{X}(t_\text{c} - W_\text{h})$ is the oldest location estimate in the look-back window.

The average speed $\bar{\mathbf{v}}$ is determined from the average distance between the projections of the historical data points in matrix $\mathbf{B}$ onto the direction vector:

$$\mathbf{b} = \mathbf{B}\hat{\mathbf{w}} \tag{7.21}$$

$$\bar{\mathbf{v}} = \frac{\sum_{i=2}^{N_\text{h}} \mathbf{b}(i) - \mathbf{b}(i-1)}{N_\text{h} - 1} \tag{7.22}$$

where, in $\mathbf{B}$, each row contains an x, y coordinate pair, and N_h is the number of historical data points.

Some example results of the movement prediction are shown in section 7.7.

7.6.2 Throughput Functions

Based on the known locations of the BS and APs and a predicted movement trajectory, the mobile device can create a throughput function for each available network.

The throughput functions for the WiFi networks are constructed using the throughput model described in chapter 3. Using this model, we have calculated the maximum throughput of all IEEE 802.11a modulation schemes for different distances as depicted in Fig. 7.5. This curve is used to characterize the achievable WiFi throughput for different distances. Table 7.1 shows the model parameters used in this work.

As mentioned in section 7.4, the proposed look-ahead prediction algorithms rely on continuously differentiable formulations of the throughput function $\hat{\Omega}(t)$, which represents the expected throughput along the predicted future

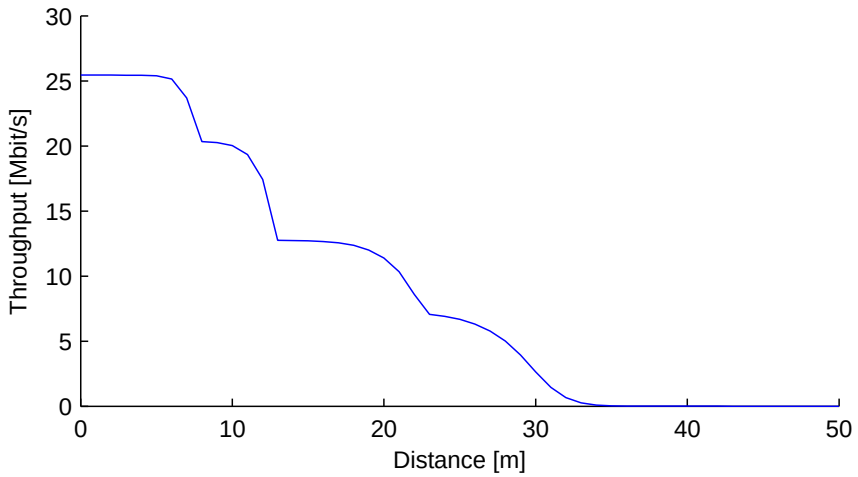


Figure 7.5: Distance-throughput relationship for WiFi networks, generated using the model in chapter 3 and the parameters in Table 7.1.

Parameter name	Values
Bit rates	$6, 9, 12, 18, 24, 36, 48, 56$ Mbit/s
Max. retransmissions ($R_{\max}$)	7
Payload size (B_{MSDU})	1024 bytes
Transmit power (P_{tx})	100 mW
Ricean K	15
Path loss exponent	2.9

Table 7.1: WiFi throughput model parameters for model described in chapter 3.

movement trajectory. As this function is dependent on the predicted trajectory relative to the APs' placements, and is therefore changing for every time the algorithms are executed, it is necessary to estimate this function efficiently for algorithm execution. For this, we have decided to use polynomial approximations for the throughput function $\hat{\Omega}(t)$. This both enables us to differentiate and find roots for the throughput functions as described in eq. (7.10) as well as estimating the throughput functions quite efficiently given a set of points along the predicted trajectory and the APs placements. The matlab function `polyfit` has been used to calculate the polynomial approximations. Since the goodness of the polynomial approximation depends on the amount of source data, which depends on W in our case, we define the order of the polynomial as:

$$M_{\text{poly}} = \min(8 + 2 \cdot \left\lfloor \frac{W}{10} \right\rfloor, 18) \tag{7.23}$$

which has been determined empirically through experiments. An example of approximated throughput functions compared to the actual throughput is shown in Figure 7.7 on page 152.

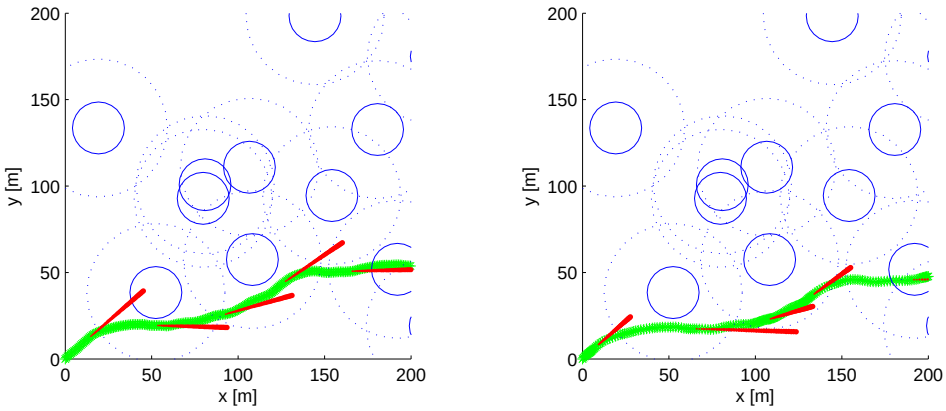
For the cellular network, we assume an urban scenario where the mobile device is able to connect using High-Speed Downlink Packet Access (HSDPA) with a bit rate of $\Omega_1 = 3.6$ Mbit/s everywhere. Of course, other cellular technologies, such as higher-rate 3rd Generation Partnership Project (3GPP) HSPA or 3GPP LTE could be used with this HO prediction scheme, as well as faster Wi-Fi technologies such as IEEE 802.11n. The proposed algorithm is not tied to specific technologies, it simply tries to exploit the situations where a local area wireless network offers higher throughput than the cellular network.

A variable throughput for the cellular network could be included in the analysis in the same way as for Wi-Fi networks. This would result in the cellular network being more often preferred at the locations where its throughput is high, and less often at the locations where its throughput is low. The constant throughput assumption is however not expected to affect the overall performance of the considered algorithms significantly.

7.7 Results and Discussion

Initially, we present two examples of how the movement prediction works for a non linear movement trajectory. This is displayed in Figure 7.6(a) and Figure 7.6(b). These results are related to the proposed handover prediction algorithms. In the shown examples the mobility prediction is performed with 20 seconds intervals and for each execution, the trajectory of the following 20 seconds is predicted. Since the user's movement trajectory is not linear, the prediction is in some cases deviating from the actual trajectory. The figures also show the impact of speed variations by the different lengths of the predicted trajectories in Figure 7.6(b).

The scenarios that are considered for evaluation of the proposed look-ahead HO prediction algorithm are specific instances of the large scale scenario presented in Figure 1.1 in chapter 1. Two different scenarios, labeled *ideal* and *realistic* have been considered as shown in Table 7.2. The *ideal* scenario assumes perfect location estimation and a constant speed linear mobility model. This scenario is used to see how well the algorithm itself performs, and how it reacts to different scenario parameter variations. Additionally, we consider a *realistic* scenario with localization errors and a random mobility model. The mobility model that is considered in this analysis is the wrap-around mobility



(a) Constant user speed (b) Varying user speed

Figure 7.6: Examples of linear movement prediction. The green line is the user's movement trajectory and the predicted trajectory 20 seconds ahead in time is indicated by the red lines, which grow in thickness (creating a cone shape) along the direction of movement. The blue circles represent the deployed Wi-Fi networks.

model described in section 2.4.3 on page 35. Table 7.2 lists the parameters that have been used in this model.

7.7.1 Ideal Scenario

In this scenario we use a linear constant speed (pedestrian) mobility model and have set the localization error std. dev. to zero.

The plot in Fig. 7.7 shows an example of a prediction window and where HOs are triggered. Further, Fig. 7.8 shows an example of the achieved throughput during a simulation run. Notice how the throughput drops to zero during the duration of each HO and how the throughput bursts when in WiFi coverage.

In the following, aggregated results for multiple independent simulation runs are presented. In the plots we show the mean including the 95% confidence interval. The first result in Fig. 7.9 shows the impact on throughput of varying the window size. For small and medium length windows the optimal K-HO algorithm is best, but for the long look-ahead windows, the heuristic algorithm performs best. The reason for this is indicated in Fig. 7.10, which shows that as the window W becomes longer, the more HOs are required within W by the optimal algorithm for the best sequence. In these simulations we have limited

Parameter name	**Default Values**	
Scenario	*Ideal*	*Realistic*
Simulation time	2000 s	
Independent simulation runs (seeds)	32	
Scenario size	1 km x 1 km	
No. of APs	250	
Hand-over delay (C)	2 s	
Max. number of HOs for opt. alg. ($K_{\max}$)	3	
Prediction window size (W)	20 s	
Historical window size (W_{h})	10 s	
Hysteresis threshold (β_{hyst})	1 Mbit/s	
Heuristic handover threshold (ρ)	$2 \cdot C \cdot \Omega_1 = 14.4$ Mbit	
Throughput variation std. dev. (σ_{TP})	2 Mbit/s	
Localization error std. dev. (σ_{pos})	0 m	2 m
Max. angular acceleration $\alpha_{\max}$	0 rad/s	0.1725 rad/s
Max. acceleration ($a_{\max}$)	0 m/s^2	0.9 m/s^2
Min. speed ($v_{\min}$)	2 m/s	0.5 m/s
Max. speed ($v_{\max}$)	2 m/s	3.5 m/s
Cellular throughput function (Ω_1)	3.6 Mbit/s	

Table 7.2: Evaluation parameters.

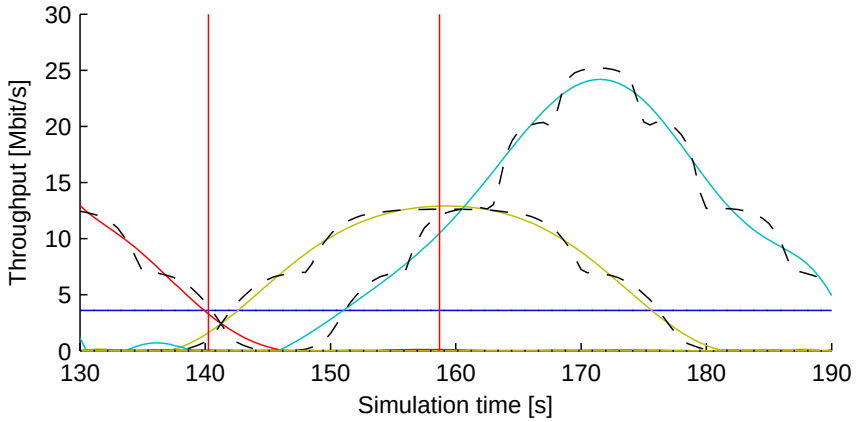


Figure 7.7: Example of 60 s prediction window for optimal algorithm. Solid curves are the approximated throughput functions of the different networks, dotted lines are actual throughput curves, and vertical lines show where HOs are performed. This view maps to the $130 - 190$ s span in Fig. 7.8.

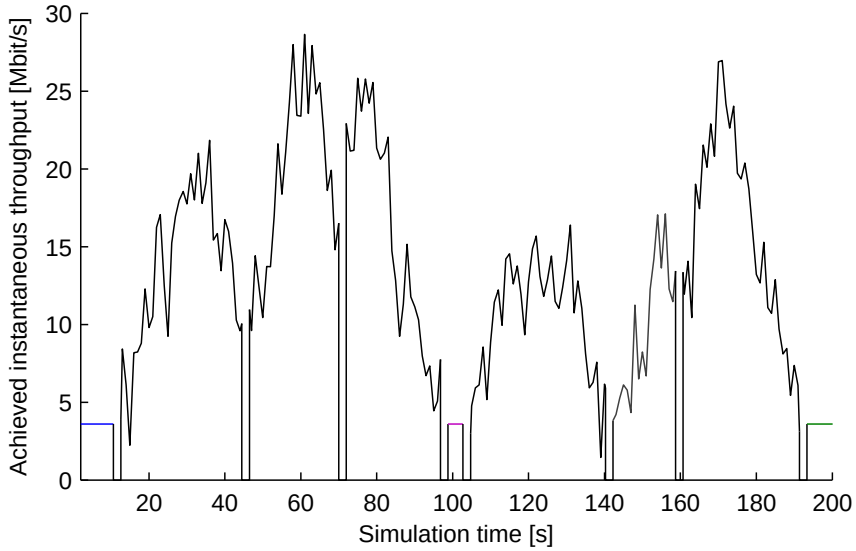


Figure 7.8: Example of the achieved throughput for a simulation run for the optimal algorithm.

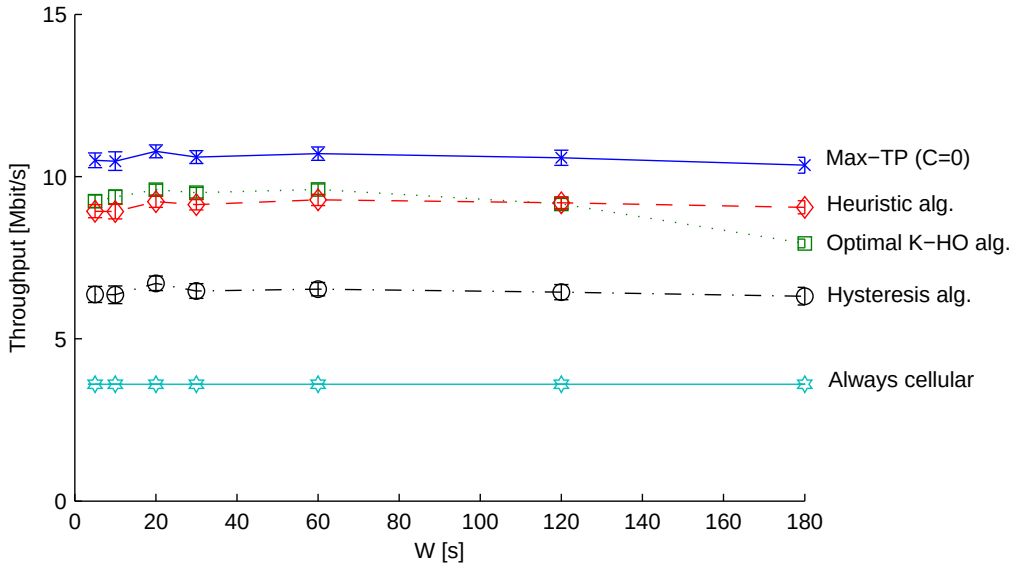


Figure 7.9: Performance impact of increasing window size (W). The legend presented in this figure applies to the following figures in this section.

the maximum allowed no. of HOs to 3 to make the simulations computationally feasible. However, since 3 HOs are highly preferred for window lengths of 120s and 180s we expect that 4 or more HOs would actually yield better results in these cases, hence the limit of maximum 3 HOs within W causes the drop in Fig. 7.9. This is supported by Fig. 7.11, which shows the average number of HOs made by the heuristic algorithm. Here it is clearly shown that the avg. number of required HOs grows asymptotically linearly with the window length, approximately for $W > 30$ seconds. However, Figure 7.9 further reveals that using a prediction window length longer than 20 s does not give any significant throughput benefit. Therefore, the limit on the number of HOs has no practical impact in this scenario.

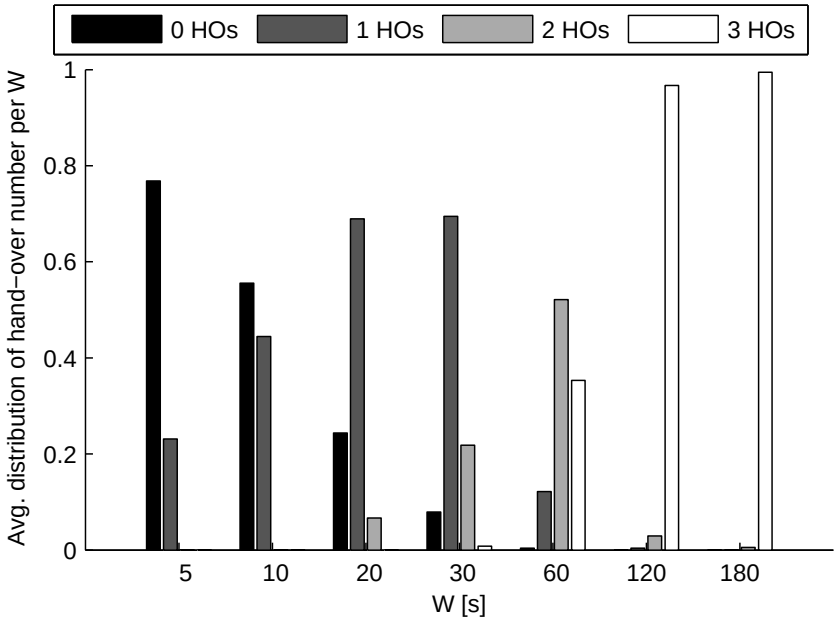


Figure 7.10: Distribution of no. of HOs for optimal algorithm for different window sizes.

In Fig. 7.12 we show how performance is improved when the density of WiFi networks increases. Notice how the optimal algorithm gains more Mbit/s than both the heuristic algorithm and the hysteresis-based algorithm when increasing the number of access points from 50 to 500. The optimal algorithm is clearly better at choosing the networks to handover to, when many options are available, even though it is significantly below the maximum throughput algorithm.

Fig. 7.13 shows the effect of increasing cost of a HO, expressed as the HO delay C. As expected the increase of C leads to a decrease in throughput.

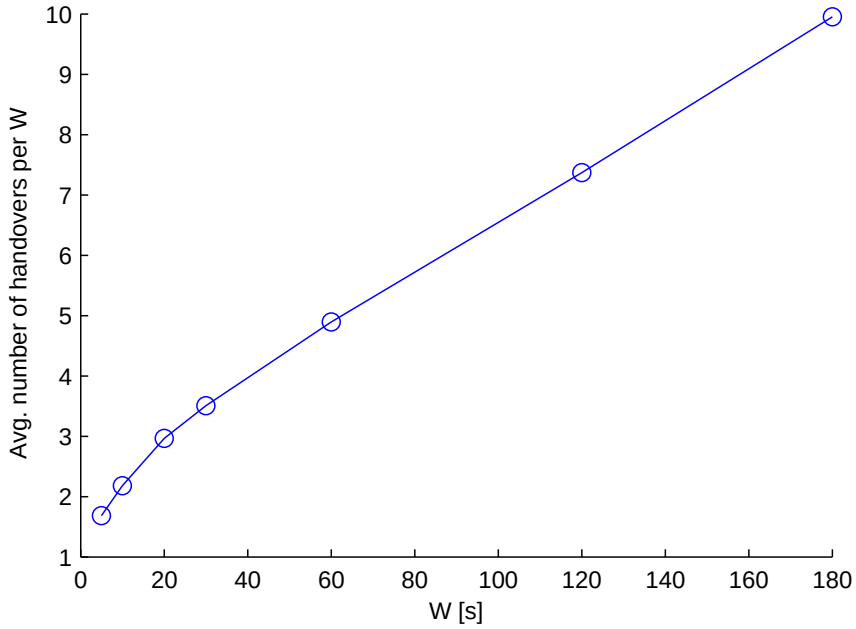


Figure 7.11: Avg. no. of HOs for heuristic algorithm for different window sizes.

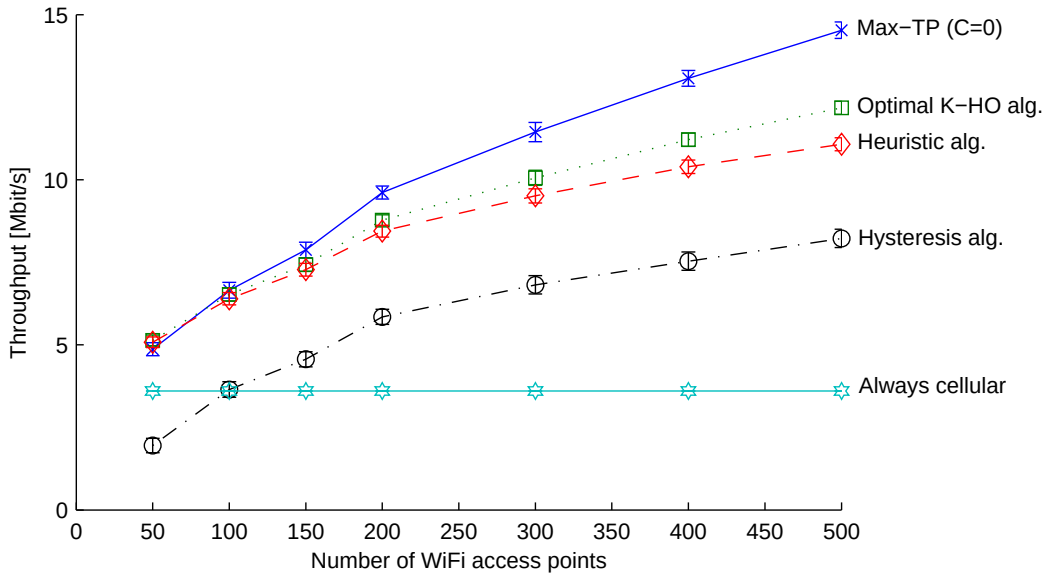


Figure 7.12: Performance impact of increasing the number of access points.

Contrary to the look-ahead algorithms, the greedy hysteresis-based algorithm suffers greatly for even small values $(0.5 - 1$ s) of C. The heuristic algorithm is gradually becoming worse than the optimal algorithm as the HO cost is increasing. This is due the heuristic algorithm not taking into account the handover delay C when determining the handover decision, as the optimal algorithm does. The maximum throughput algorithm is of course not affected by the increasing C, since it assumes that C is always zero.

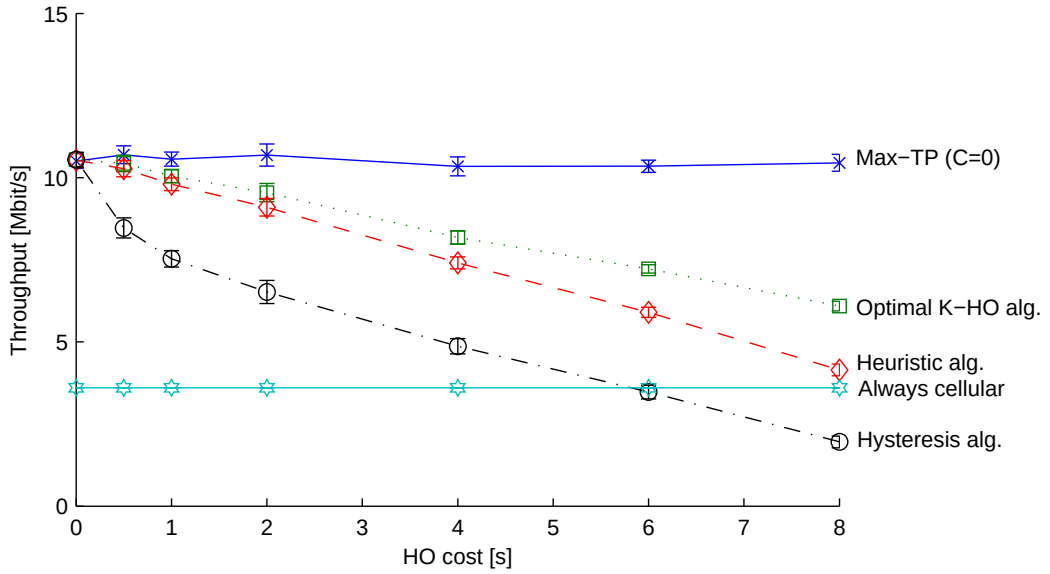


Figure 7.13: Performance impact of increasing handover delay C.

The effect of increasing the localization error is shown in Fig. 7.14. The plot clearly shows, that performance deteriorates with higher localization inaccuracy, since it leads to erroneous movement prediction and in turn bad handovers.

Fig. 7.15 shows how the hysteresis-based algorithm performs worse the more the instantaneous throughput varies. Since the look-ahead algorithm uses a priori knowledge of average throughput levels and the actual throughput varies around the mean, it is not affected similarly.

Finally, Fig. 7.16 shows the effect of varying the maximum angular acceleration $\alpha_{\max}$. Larger values of $\alpha_{\max}$ reduces the accuracy of the movement prediction, thus the achieved throughput drops for the prediction algorithms.

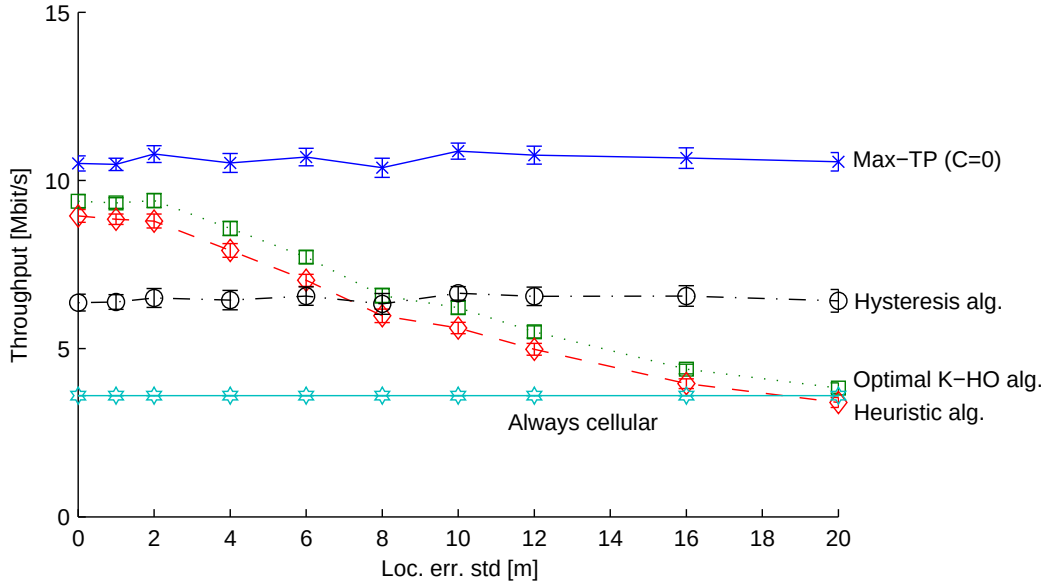


Figure 7.14: Performance impact of increasing localization error.

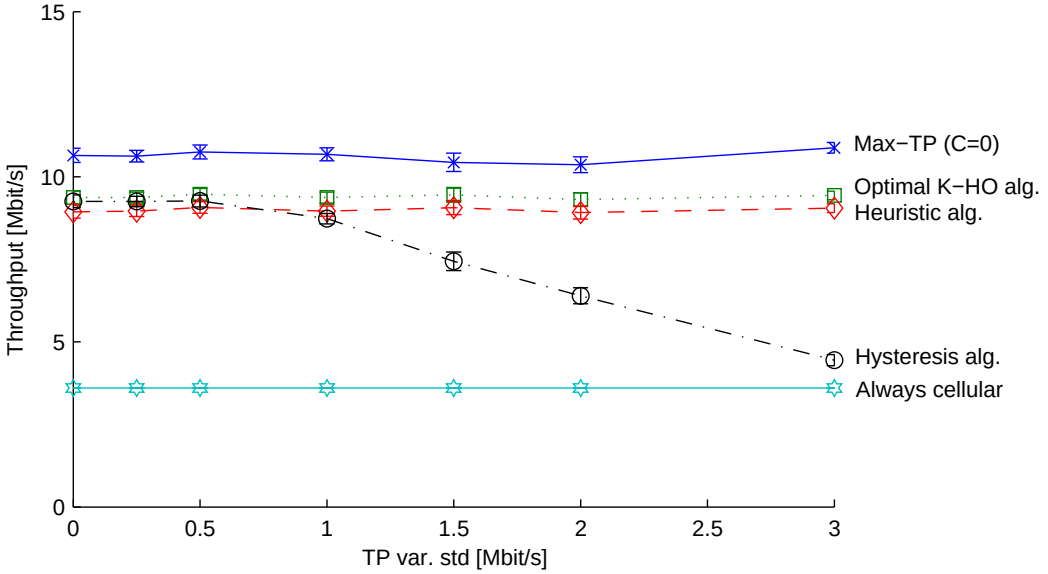


Figure 7.15: Performance impact of increasing throughput variation.

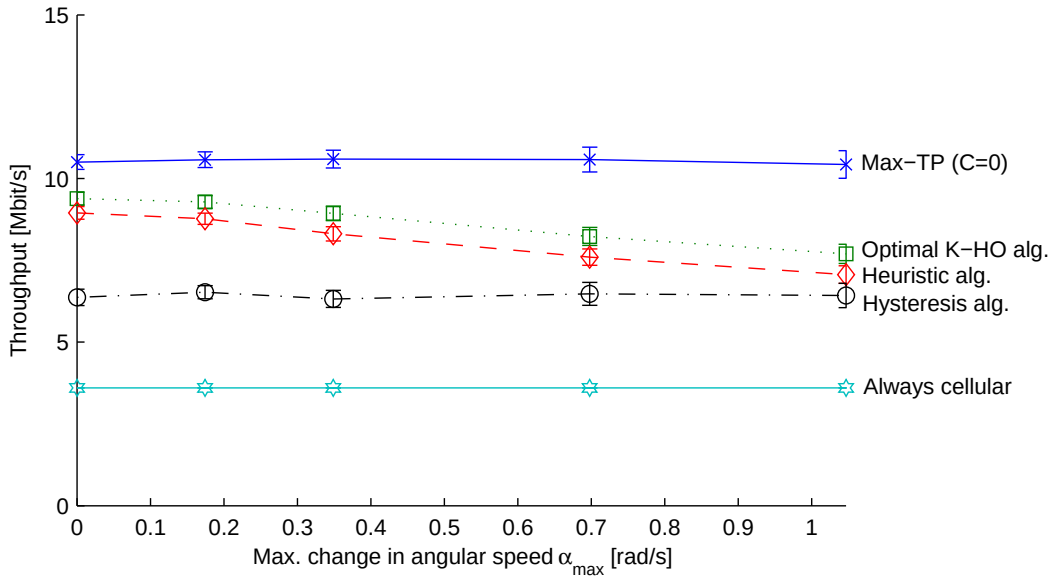


Figure 7.16: Performance impact of increasing $\alpha_{\max}$.

7.7.2 Realistic Scenario

Considering now the scenario with more realistic mobility model parameters and location error, given by the parameters in the right-most column in Table 7.2, we show in Fig. 7.17 how the algorithms are affected by different look-ahead window sizes. Now, as the window size increases, the localization errors and the random mobility model that are considered in the realistic scenario result in a more rapid decrease in throughput. This means, that while a long look-ahead window where several HOs are planned looked usable in an ideal system, then in practical systems with imperfect movement prediction, shorter prediction windows are necessary. Only in cases where the movement prediction is good, longer look-ahead windows are useful. These observations further motivate an improved prediction scheme, where instead of planning W seconds ahead and then waiting until the prediction window has elapsed before running the prediction algorithm again, it would make sense to run the prediction algorithm more often and then update the planned handover sequence correspondingly. Also, weights could be introduced in the cost function to make handovers far ahead in time less significant, as a reflection of the movement prediction uncertainty.

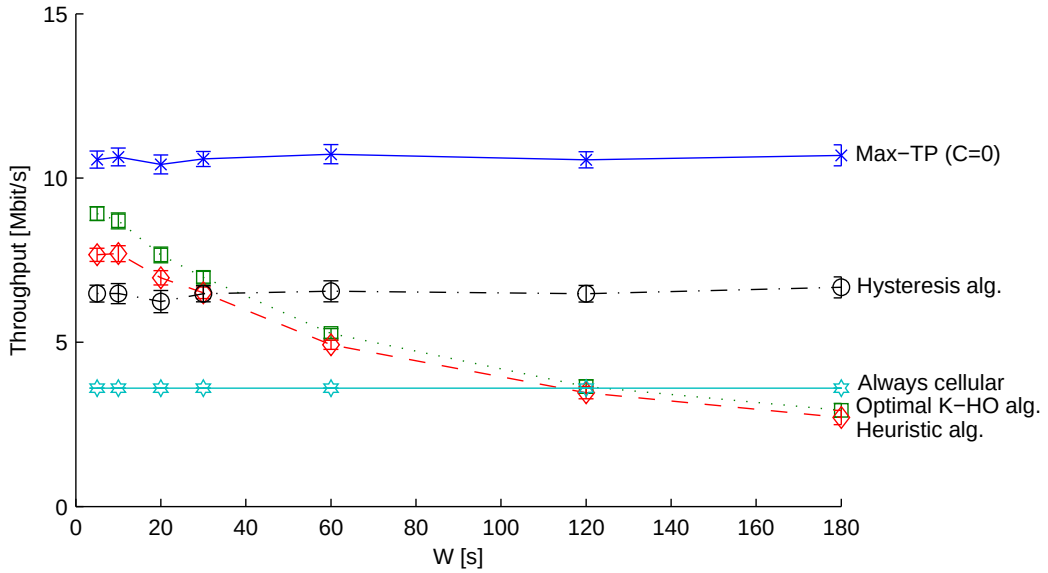


Figure 7.17: Performance impact of increasing window size (W), in realistic scenario.

7.8 Implementation Considerations

The nature of the K-HO optimal and heuristic look-ahead prediction algorithms do not dictate that they need to be implemented in the mobile device or in the network. Also, the MIH framework in principle supports both device and network triggered handovers, as described in [Baek et al., 2008]. However, the algorithms have some dependencies that make the network-based approach most attractive:

Processing power First and foremost we have noticed in our simulation prototype that the processing power required to determine the best HO sequence with the optimal algorithm is quite substantial for 4 or more HOs within W. A battery-driven device may therefore experience a significant reduction in battery life-time if these calculations are performed locally. However, the heuristic algorithm requires less processing power and could therefore be implemented in the mobile device.

Database look-up Secondly, the algorithms need to look up the average throughput of networks along the mobile device's expected future trajectory.

Doing these regular database look-ups over a wireless link, could incur a significant increase in the signaling overhead. A solution could be for the mobile device to maintain a cache of this information.

Resource sharing In practical networks it may be necessary to account for the number of collocated users and their instantaneous load on the different APs and BSs, as users need to share resources. This advocates for a network-based approach, since the information about other users is available on the network side.

Location updates Finally, the algorithms rely on a prediction of the mobile device's future movement trajectory. In cases where this is based on only GPS location estimates or a distributed localization algorithm the device-based solution could be attractive. However, if a centralized network-based approach where fusion of measurements from e.g. GPS and cellular and WiFi networks is considered, as in chapter 4, then the network-based HO prediction algorithm would allow the functions to be implemented in the same physical entity and thus reduce inter-process communication latency.

7.9 Conclusion and Outlook

In this work we have considered the problem of determining when to handover and which network(s) to handover to, within a fixed look-ahead window for a multi-network scenario, in order to maximize the achieved throughput of a mobile multi-radio terminal. Based on an analytical formulation of the handover problem, we have proposed an optimal and a heuristic algorithm for this and compared them to a simple hysteresis-based algorithm and the case where the cellular network is always used. The optimal algorithm finds the optimal handover sequence for up to K handovers within the look-ahead window. In this work we have found that the optimal algorithm is computationally feasible for up to 3-4 handovers, whereas the proposed heuristic algorithm works with any practical number of handovers. The algorithms have been implemented and evaluated using simulations in matlab for a scenario with ubiquitous cellular coverage and randomly scattered high-speed WiFi hotspots.

Our results for the ideal scenario where the movement prediction is assumed to be perfect, have shown that the optimal algorithm achieves the highest throughput for cases where less than 4 handovers are required. For longer

look-ahead windows where more handovers are needed, the heuristic algorithm achieves the highest throughput.

The two look-ahead algorithms are equally affected by localization errors, where errors of up to appr. 4 m std. dev. result in only a minor drop in performance. Comparing this to the results from Chapter 4 where the standard deviation is in the order of 2-3 m, these results show that the achievable localization accuracy is sufficient for location based handover optimization. However, for the GNSS and cellular based localization systems presented in Figure 2.5 on page 27 the estimated standard deviations of 8-13 meters would mean that the hysteresis based algorithm would be superior to both of the prediction algorithm.

A general prerequisite for the look-ahead prediction algorithms is an accurate movement prediction. Our results for a realistic scenario shows that inaccurate movement prediction strongly limits the look-ahead window length. However, in cases where the movement of the user is constrained physically by e.g. roads, sidewalks or walls, this can be exploited for improved movement prediction.

For this work, all handovers have been assumed to be hard, "break-before-make" handovers. For handovers between cellular and Wi-Fi networks, the use of Media Independent Handover (MIH) and Mobile IP (MIP) technologies can deliver seamless handovers if the user terminal has a dedicated radio for each technology. However, for handovers between two Wi-Fi networks that are not part of the same Basic Service Set (BSS) or Extended Service Set (ESS), the terminal would need two radios for Wi-Fi in order to associate and prepare the handover to another network while still having an ongoing connection at the first network. As two Wi-Fi radios are uncommon in user terminals, it is assumed that a break-before-make handover is necessary. Even though handovers between Wi-Fi and cellular are seamless, the handover still incurs a signaling overhead, which has a cost for the operator. For simplicity in the analysis, the cost of any handover was represented as a handover delay, since this would keep the number of handovers low. This representation has however most likely exaggerated the cost of seamless handovers between cellular and Wi-Fi in the analysis, which means that the achieved throughput is lower than it should be. Considering that the goodness of the handover algorithms has been measured only in terms of throughput and not signaling overhead, it is however useful that seamless handovers are not cost-free but are punished in terms of throughput instead. A possibly extension of this work would be to consider a cost function based approach which takes into account both delay and signaling overhead for different types of handovers.

An obvious future work item is to reconsider how the look-ahead predictions are used. In this work, the handover decisions made within one prediction window is not re-evaluated as time passes. However, a better handover sequence may be achievable if for example a new prediction window is made after each performed handover, or with regular intervals shorter than the prediction window. In this way the impact of inaccuracies in movement prediction is kept low. Also, weights could be introduced in the cost function to reflect the movement prediction uncertainty and thereby give less significance to decisions far ahead in time.

In this work a fingerprinting/throughput database has been assumed to be available for a priori estimation of achievable throughput for the available networks. For a practical system it would be necessary to investigate the accuracy of the predictions coming from such a database, as this is a crucial factor for good system performance. The database could be populated either from ray-tracing simulations, from dedicated measurement campaigns, or by letting users' mobile terminals submit information of observed networks and channel characteristics to a central database.

In this work the handover decision has not taken into account the available resources at APs and BSs. However, in actual networks, the available resources are typically shared between users and it would therefore make sense to take this aspect into account. For the case of a single mobile user, where the load imposed by other users is relatively constant, the proposed selection algorithms can be used as long as the expected throughput along the predicted movement trajectory can be retrieved from the database. In the case where multiple users are performing handovers - potentially to and from the same networks, it is necessary to extend the optimization problem to consider multiple users jointly, since the decision taken by one user will affect the decisions of the remaining users. Given the computational effort required to find the optimal solution for the single user case, it does not seem immediately feasible to consider the optimal solution for the multi-user case with the proposed method. Here, it would be necessary to consider more efficient algorithms, such as cost function based algorithms, e.g., [Sun et al., 2008, Zhang et al., 2006] or combination/heuristic algorithms, e.g., [Xia et al., 2007].

References

J.Y. Baek, D.J. Kim, Y.J. Suh, E.S. Hwang, and Y.D. Chung. Network-initiated handover based on IEEE 802.21 framework for QoS service continuity in UMTS/802.16 e networks. In *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, pages 2157–2161. IEEE, 2008.

S. Balasubramaniam and J. Indulska. Vertical handover supporting pervasive computing in future wireless networks. *Computer Communications*, 27(8): 708–719, 2004.

W.T. Chen and Y.Y. Shu. Active application oriented vertical handoff in next-generation wireless networks. In *2005 IEEE Wireless Communications and Networking Conference*, pages 1383–1388, 2005.

Cisco Systems. Global Mobile Data Traffic Forecast Update, 2009-2014. *Cisco Systems Feb 9th*, 2010.

P. Deshpande, X. Hou, and S.R. Das. Performance Comparison of 3G and Metro-Scale WiFi for Vehicular Network Access. *10th ACM Internet Measurement Conference (IMC 2010), November 1-3*, 2010.

Richard Gass and Christophe Diot. An experimental performance comparison of 3g and wi-fi. In *Passive and Active Measurement*, volume 6032 of *Lecture Notes in Computer Science*, pages 71–80. Springer, 2010.

IEEE. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999)*, pages C1–1184, 12 2007.

I.T. Jolliffe. *Principal component analysis*. Springer series in statistics. Springer, 2002. ISBN 9780387954424.

M. Kassar, B. Kervella, and G. Pujolle. An overview of vertical handover decision strategies in heterogeneous wireless networks. *Computer Communications*, 31(10):2607–2620, 2008.

C.W. Lee, L.M. Chen, M.C. Chen, and Y.S. Sun. A framework of handoffs in wireless overlay networks based on mobile IPV6. *Selected Areas in Communications, IEEE Journal on*, 23(11):2118–2128, 2005. ISSN 0733-8716.

S. Mohanty and I.F. Akyildiz. A cross-layer (layer 2+ 3) handoff management protocol for next-generation wireless systems. *IEEE Transactions on Mobile Computing*, pages 1347–1360, 2006. ISSN 1536-1233.

K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J.P. Makela, R. Pichna, and J. Vallstron. Handoff in hybrid mobile data networks. *IEEE Personal Communications*, 7(2):34–47, 2000.

K. Ramachandran, S. Rangarajan, and J.C. Lin. Make-before-break mac layer handoff in 802.11 wireless networks. In *Communications, 2006. ICC'06. IEEE International Conference on*, volume 10, pages 4818–4823. IEEE, 2006. ISBN 1424403553.

E. Simoncelli. Least squares optimization. *Lecture Notes, http://www. cns. nyu. edu/eero/teaching. html*, 2003.

C. Sun, E. Stevens-Navarro, and V. Wong. A Constrained MDP-based Vertical Handoff Decision Algorithm for 4G Wireless Networks. In *Conference on Communications*, pages 2169–2174. Citeseer, 2008.

L. Xia, L. Jiang, and C. He. A novel fuzzy logic vertical handoff algorithm with aid of differential prediction and pre-decision method. In *Communications, 2007. ICC'07. IEEE International Conference on*, pages 5665–5670. IEEE, 2007. ISBN 1424403537.

X. Yan, N. Mani, and YA Cekercioglu. A traveling distance prediction based method to minimize unnecessary handovers from cellular networks to WLANs. *Communications Letters, IEEE*, 12(1):14–16, 2008. ISSN 1089-7798.

X. Yan, Y. Ahmet Sekercioglu, and S. Narayanan. A survey of vertical handover decision algorithms in Fourth Generation heterogeneous wireless networks. *Computer Networks*, 54(11):1848–1863, 2010. ISSN 1389-1286.

K. Yang, I. Gondal, B. Qiu, and L.S. Dooley. Combined SINR based vertical handoff algorithm for next generation heterogeneous wireless networks. In *Global Telecommunications Conference, 2007. GLOBECOM'07. IEEE*, pages 4483–4487. IEEE, 2007.

A.H. Zahran, B. Liang, and A. Saleh. Signal threshold adaptation for vertical handoff in heterogeneous wireless networks. *Mobile Networks and Applications*, 11(4):625–640, 2006. ISSN 1383-469X.

J. Zhang, HC Chan, and V. Leung. A location-based vertical handoff decision algorithm for heterogeneous mobile networks. *Proc. of IEEE GLOBECOM*, 2006.

8

Conclusions and Outlook

The increasing popularity of mobile computing platforms such as smartphones and tablets, challenges mobile operators' cellular networks with an expected annual doubling of mobile data amounts. This motivates the use of techniques that can offload the cellular networks using locally available networks, such as Wi-Fi. Further, as location information is becoming ubiquitously available through network based localization techniques and cooperation, the main problem of this thesis has been:

How well suited are location based network optimizations for wireless last hop route selection in networks with mobile users?

In this thesis, location based network optimizations were considered for the following two scenarios:

- The large scale scenario has ubiquitous cellular coverage and scattered Wi-Fi hotspots. Here, the focus was on location based handover algorithms that allows a mobile user to use Wi-Fi networks when available. Further, a large scale scenario was used for the investigation of the achievable location accuracy.
- The small scale scenario concentrated on a single Wi-Fi access point, and the users associated to it. Specifically, location based relaying techniques were considered for this scenario.

The central focus area in this work was to compare traditional measurement based approaches to location based approaches, taking into account the errors resulting from provisioning if input information and the errors that relates to the different properties of these approaches. Notice that the overall objective of this work has not been to propose relaying and handover protocols that are better than existing state of the art protocols, but rather to study the benefits and drawbacks of using location information compared to traditional schemes using link quality measurements. Table 8.1 presents an overview of the types of errors that were considered for the five different contribution chapters in the thesis.

Distance-dependent Throughput Model for Wi-Fi This throughput model describes the expected achievable throughput between a transmitter and receiver in a Wi-Fi network, based on the path loss between the nodes. The contribution consists of a formulation of the expected link throughput as a function of received power, interference, bit error rate and frame error rate. The model accounts for the impact of frame errors on throughput, by modeling the backoff mechanism in IEEE 802.11 DCF. The model does not take into account that neighboring users may cause collisions, since the model was developed for a priori estimation of link performance for the location based network optimizations considered in chapters 5, 6, and 7. In addition to using the model for a priori link estimation, the model has also been used for evaluation of throughput performance in those chapters. The throughput model has been considered sufficient for evalution, as the important aspects of these analyses are to gain first insights on the impact of different error terms and to establish basic relations between measurement and location based network optimizations.

Error types	**Ch3**	**Ch4**	**Ch5**	**Ch6**	**Ch7**
node mobility		X	X		X
measurement collection delay		X	X		
small-scale fading	X		X (SNR var, BER calc)	X (BER calc)	X (TP var, BER calc)
inaccurate prop. model par.			X		
interference	X			X	
location accuracy		outcome	X	X	X
movement prediction inaccuracy					X

Table 8.1: Overview of which error types were considered in which chapters.

Realistic Communication Constraints for Localization The next contribution of this thesis was a study of how realistic communication constraints affect conventional and cooperative localization and tracking algorithms. The contribution, which is described in chapter 4, served as a means to understand the achievable location information accuracy of a realistic network based localization system. The specific contribution in this thesis for this joint work was the accurate modeling and evaluation of the message exchanges required to realize conventional and cooperative network-based localization algorithms. For the cooperative localization, a group mobility model was used to generate correlated user movements. Localization was enabled through UWB based ranging measurements between mobile user and anchor nodes as well as between mobile users for the cooperative algorithm.

The results showed that the difference between not considering the measurement collection and using a realistic model of the measurement collection was significant. For the cooperative tracking algorithm the 90 %-error went from 3 m to 4.5 m. The results also indicated that a large number of cooperating users lead to decreasing performance due to network congestion. The evaluation results showed that the tracking error distribution was close to Gaussian in some cases, and that tracking errors of just 2-3 meters standard deviation can be achieved for the cooperative EKF algorithm, if a sufficient number of anchor nodes is deployed in the considered scenario. Compared to the accuracy requirements of maximum 3-4 meters standard deviation for the relaying and handover network optimizations considered in Chapter 5 and Chapter 7, the performance of the considered cooperative localization system is considered sufficient. For the SimTX scheme the use of location information for both relay selection and power adaptation has proven to be more sensitive to location error than pure relay selection. With location errors in the order of 2-3 meters standard deviation the scheme is still superior or on par with the two-hop relaying scheme.

Location-based Relaying The contribution concerning two-hop relaying in Chapter 5, was focused on comparing the benefits and drawbacks of location based and measurement based relaying schemes, when considering the impact of node mobility, measurement collection delay, small-scale fading, inaccurate propagation model parameters, and inaccurate location information.

The study has shown that the amount of signaling overhead required to collect location measurements is linearly dependent on the number of mobile nodes, whereas the overhead for the link SNR measurement based scheme has

a quadratic relationship. This means that in practice, a higher update frequency is possible with the location based scheme, thereby letting this scheme cope better with mobility. As the location based scheme relies only on a path loss model for performance prediction, the performance of the scheme is likely dependent on the accuracy of the model parameters. Results from the study showed that inaccuracies of the path loss exponent within ± 1.5 of the true value lead to near optimal results for a BER relay selection criteria, whereas a throughput criteria had an acceptable range of only ± 0.3. Identifying NLOS cases was found to be critical for the location based scheme. Here, binary LOS/NLOS knowledge was found to yield close to optimal results, if the NLOS attenuation parameter was within ± 5 dB of the true value. The provision of such information however requires detailed knowledge of the scenario geometry as for example where walls are located. In summary, relaying based on location information is more efficient, but requires infrastructure support and environment knowledge. It would be useful to combine with environment self-learning approaches such as SLAM. Another future work item would be to develop an analytic model which is able to judge the usefulness of different relay choices when taking into account the possible movements, localization inaccuracy, and measurement age. This could be realized by extending the work in [Olsen et al., 2010] to include the relaying aspects.

Simultaneous Transmissions in Relaying The contribution described in Chapter 6, focused on an extension of the location based relaying scheme that allows simultaneous relay-to-destination transmissions through interference-aware power adaptation called SimTX. By allowing simultaneous transmissions the SimTX scheme was shown to increase the overall downlink throughput. The efficiency of the scheme depends on the availability of suited pair destination nodes, which is more likely in dense scenarios. In such cases the utilization of the AP may already be high, and here the SimTX scheme helps to further increase the capacity. The impact of inaccurate location information was studied, and the results of the study showed that the SimTX scheme was more sensitive to location inaccuracies than the sequential two-hop relaying scheme. This means that while the SimTX scheme brings a significant performance improvement under ideal conditions, its higher sensitivity to location inaccuracy means that it should be used carefully in less ideal conditions.

The SimTX scheme has been considered for the case of having two simultaneous transmissions. In [Hu and Tham, 2010] the authors have calculated the maximum number simultaneous transmissions to be 5 for the CCMAC scheme.

Since this scheme is quite similar to the SimTX scheme, except that it considers the uplink case with simultaneous source to destination transmissions, and that it is measurement based.

In this thesis the proposed scheme has been evaluated for the example use case of IEEE 802.11a networks, however it would be similar applicable for the other common 802.11 variants b and g. The only required change would be a change of the model parameters in the throughput model, as well as the models used for estimating the BER for different modulation schemes and SINR levels.

Handover Optimization The last contribution in Chapter 7 dealt with exploiting the prediction power of location information, through movement prediction, for handover optimization in heterogeneous networks. The problem of deciding when to handover between different available networks within a fixed time horizon has been formulated as an optimization problem. The optimal solution of the handover problem is outlined, assuming continuous differentiability of the functions used to describe the expected throughput. This solution was however not useful for an online decision algorithm, if more than 3-4 handovers within the prediction window were considered. As an alternative, a heuristic algorithm, which is feasible for online use, was proposed. The location based algorithms were evaluated numerically and compared to some selected reference schemes, in order to determine the impact of node mobility, location information inaccuracy, and movement prediction inaccuracy.

The proposed optimal algorithm, which was only allowed up to four handovers within the prediction window for feasibility reasons, was found to outperform the heuristic algorithm, especially for scenarios with high densities of Wi-Fi networks. However, for longer prediction windows the heuristic algorithm, which was not limited in the number of handovers, achieved better throughput than the limited optimal algorithm. In practice, long prediction windows may have limited use, since very accurate movement prediction is required for good performance.

Like the location based relaying algorithm, the proposed location based handover algorithms rely on a priori environment knowledge in the form of estimates of achievable throughput along the predicted movement trajectory for the different available networks. For accurate movement prediction, the handover prediction algorithm were found to work well. However, in cases with imperfect movement prediction, the handovers that were planned in the end of the prediction window were not good choices in practice. A possible way to overcome this, is to have a sliding prediction window, which frequently reevaluates the choices made previously.

Furthermore, for the approach to work in actual networks where multiple users are sharing resources, an important extension would be to consider the load imposed by other users, as a throughput estimate that is based solely on path loss, does not represent the achievable throughput of a highly loaded access point very well.

Benefits of Location Information Finally, a summary is given over the key benefits of using location based network optimizations for wireless last hop route selection in networks with mobile users.

- Collection of location information requires less signaling overhead than link measurements.
- Exploitation of geometrical relations was found to be useful for performance prediction of different network configurations for relay selection, by accounting for path loss between nodes. In cases where obstructions, such as walls, cause shadow fading, accurate predictions require a priori knowledge of the environment in order to estimate the link qualities.
- Geometrical relations given by location information was also exploited for a priori interference prediction, thus allowing coordination of simultaneous transmissions in mobile networks. This however required slightly higher location accuracy than the location based relay selection. Traditional measurement based approaches for interference adaptation, adapt to interference over time, and are therefore not suited for mobile networks.
- Tracking the location of nodes over time allows for movement prediction, which can be exploited to predict future link conditions and connectivity options. This was found to be useful for handover optimization. For cases where movement prediction is inaccurate, a sliding prediction window and the introduction of cost function weights has been identified as useful extensions.

The contributions in this thesis have clearly shown that location information is useful for network optimizations. Especially the knowledge of geographical relations between network entities given from the location information and the possibilities for proactivity that movement prediction gives, are believed to be essential features in future network optimizations. These features make it possible for network optimizations to work efficiently in complex and highly mobile scenarios. A promising exploitation of location information would be to consider hybrid approaches, e.g., for relaying, where location information

is used for providing initial guesses of the optimal configuration, whereafter a set of selected measurements are obtained for fine tuning. In a prediction context, e.g., for handover optimization, location information could be used for planning ahead and again selected measurements could be used to precisely trigger the network reconfiguration.

References

Zhengqing Hu and Chen-Khong Tham. Ccmac: Coordinated cooperative mac for wireless lans. *Computer Networks*, 54(4):618 – 630, 2010. ISSN 1389-1286. doi: 10.1016/j.comnet.2010.02.001. Advances in Wireless and Mobile Networks.

R.L. Olsen, J. Figueiras, J.G. Rasmussen, and H.P. Schwefel. How precise should localization be? A quantitative analysis of the impact of the delay and mobility on reliability of location information. *Proc. of IEEE GLOBECOM*, 2010.

A

Timing Specifications of IEEE 802.11a

This appendix describes the constants and the calculation of the different variables used for the throughput model described in Chapter 3. The given numbers and equations relate to IEEE 802.11a DCF in basic (acknowledged) mode and are based on the protocol specifications in [IEEE, 2007].

Parameter	**Value**
T_{SIFS}	9 μs
T_{DIFS}	34 μs
N_{ACK}	112 bits

Table A.1: Constants for 802.11a MAC DCF, Basic mode.

The following variables are used in the calculations for the throughput model in chapter 3.

$$\overline{T_{\mathrm{BO}}}(r) = \frac{\min(1023, 2^{(4+r)} - 1)}{2}\ \mu s \tag{A.1}$$

$$T_{\mathrm{data}} = 16 + 4 + 4 \left\lceil \frac{8 \cdot (B_{\mathrm{MPDU}}) + 1 + \frac{6}{R_{\mathrm{code}}}}{N_{\mathrm{DBPS}}} \right\rceil\ \mu s \tag{A.2}$$

$$T_{\mathrm{ACK}} = 16 + 4 + 4 \left\lceil \frac{8 \cdot 14 + 1 + \frac{6}{R_{\mathrm{code}}}}{N_{\mathrm{DBPS}}} \right\rceil\ \mu s \tag{A.3}$$

$$N_{\mathrm{data}} = (B_{\mathrm{MPDU}}) \cdot 8 \text{ bits} \tag{A.4}$$

$$N_{\mathrm{ACK}} = (2 + 2 + 6 + 4) \cdot 8 \text{ bits} \tag{A.5}$$

where r is the retry count, B_{MPDU} is the MAC frame size in octets, N_{DBPS} is the number of data bits per OFDM symbol, and R_{code} is the coding rate. N_{DBPS} and R_{code} depend on the used modulation scheme and are specified in Table A.2.

Scheme	$\mathbf{R}_{\mathrm{code}}$	$\mathbf{N}_{\mathrm{DBPS}}$
6 Mbit/s	1/2	24
9 Mbit/s	3/4	36
12 Mbit/s	1/2	48
18 Mbit/s	3/4	72
24 Mbit/s	1/2	96
36 Mbit/s	3/4	144
48 Mbit/s	2/3	192
54 Mbit/s	3/4	216

Table A.2: Code rate and number of bits per OFDM symbol for the different 802.11a PHY modes.

For the throughput model, which is used in Chapter 3, 5, 6, and 7, we use $B_{\mathrm{MPDU}} = 36 + B_{\mathrm{MSDU}}$, which accounts for MAC data frame header (28 octets) and LLC+SNAP headers (3+5 octets). In Chapter 4, the measurement frames are MAC data frames, meaning that $B_{\mathrm{MPDU}} = 28 + B_{\mathrm{MSDU}}$. Finally, the measurement frames in Chapter 5 are assumed to be MAC control frames, meaning that $B_{\mathrm{MPDU}} = 20 + B_{\mathrm{MSDU}}$.

References

IEEE. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999)*, pages C1–1184, 12 2007.